



WebMedia2018

XXIV SIMPÓSIO BRASILEIRO DE SISTEMAS MULTIMÍDIA E WEB

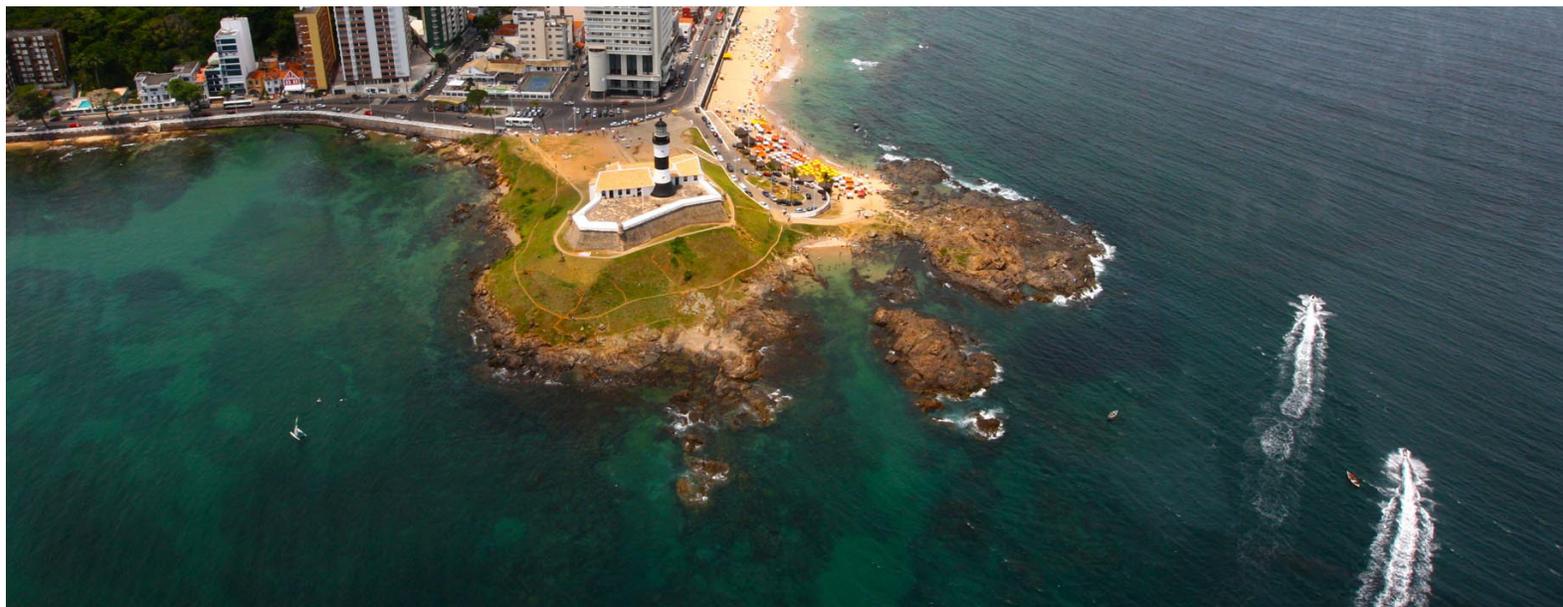
Realização



16 a 19

Outubro

Salvador - BA - Brasil



Workshops

Organizadores

Manoel C. Neto (IFBA)
Renato L. Novais (IFBA)
Eduardo Barrére (UFJF)
Frederico Durão (UFBA)
Carlos de Salles (UFMA)
Álan Lívio (PUC-Rio)
Tiago Maritan (UFPB)
Antonio Carlos Souza (IFBA)
Antonio M. Pitangueira (IFBA)
Leandro N. Ciuffo (RNP)
Valter Roesler (UFRGS)
Alexandre Carissimi (UFRGS)

Organização



Cooperação



Patrocínio



Editora

Sociedade Brasileira de Computação (SBC)



WebMedia2018

XXIV SIMPÓSIO BRASILEIRO DE SISTEMAS MULTIMÍDIA E WEB

XXIV Simpósio Brasileiro de Sistemas Multimídia e Web

De 16 a 19 de Outubro de 2018

Salvador, Bahia, Brasil

ANAIS

Workshops e Pôsteres

Organizadores

Manoel C.M. Neto (IFBA)

Renato L. Novais (IFBA)

Eduardo Barrére (UFJF)

Frederico Durão (UFBA)

Carlos de Salles (UFMA)

Álan Lívio (PUC-Rio)

Tiago Maritan (UFPB)

Antonio Carlos Souza (IFBA)

Antonio M. Pitangueira (IFBA)

Leandro N. Ciuffo (RNP)

Valter Roesler (UFRGS)

Alexandre Carissimi (UFRGS)

Realização

Sociedade Brasileira de Computação – SBC

Em cooperação com

ACM/SIGWEB e ACM/SIGMM

Organização

Instituto Federal da Bahia – IFBA

Biblioteca Raul V. Seixas – Instituto Federal de Educação, Ciência e Tecnologia da Bahia - IFBA - Salvador/BA.

Responsável pela catalogação na fonte: Samuel dos Santos Araújo - CRB 5/1426.

S613a Simpósio Brasileiro de Sistemas Multimídia e Web. Workshops e Posterres (24. : 2018 : Salvador, BA).

Anais [recurso eletrônico] / organização: IFBA. Salvador: SBC; IFBA, 2018.

183 p.

E-book.

ISBN: 978-85-7669-435-9.

Evento promovido pela Sociedade Brasileira de Computação – SBC, Porto Alegre, RS, em 16 a 19 de outubro, 2018, Salvador, Bahia. Organização: Instituto Federal de Educação, Ciência e Tecnologia da Bahia – IFBA.

1. Multimídia. 2. Web. 3. Redes Sociais. 4. Computação Ubíqua e Móvel. I. SBC. II. IFBA. III. Título.

CDU 2 ed. 004.7

Prefácio

O WebMedia é um evento anual, realizado pela Sociedade Brasileira de Computação (SBC) e, desde 1995, constitui-se no mais importante fórum no Brasil para debates entre pesquisadores, profissionais e estudantes das áreas de Multimídia e Web.

A vigésima quarta edição do WebMedia ocorre no período de 16 a 19 de outubro de 2018, na cidade de Salvador, Bahia. A organização do WebMedia 2018 está sob a responsabilidade do Instituto Federal da Bahia (Bahia). O simpósio ocorre com patrocínio do Núcleo de Informação e Coordenação do Ponto BR (NIC.br), Comitê Gestor da Internet no Brasil (CGI.br), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Os anais do WebMedia 2018 contendo os artigos completos e curtos aceitos nesta edição estão publicados na ACM Digital Library. Neste documento, que corresponde aos anais estendidos do WebMedia 2018, estão publicados os artigos dos workshops organizados no escopo do evento, a saber: o XVIII Workshop de Teses e Dissertações (WTD), o XVII Workshop de Ferramentas e Aplicações (WFA) e o XV Workshop de Trabalhos de Iniciação Científica (WTIC). Além dos artigos dos workshops do WebMedia 2018, este volume estendido contém os anais do V workshop “O Futuro da Videocolaboração”, coordenado pelo CT-Video (Comitê Técnico de Prospecção Tecnológica em Videocolaboração).

Deve ser salientado o enorme esforço dos membros dos Comitês de Programa do WebMedia 2018 e dos seus workshops associados para realizar a seleção dos trabalhos a serem apresentados, dentre tantas submissões de excelente qualidade. Agradecemos a todos os autores que submeteram seus artigos e congratulamos aqueles que tiveram seus trabalhos aceitos. Por último, gostaríamos de agradecer aos membros dos comitês de organização e de programa, bem como aos revisores e a todos aqueles que voluntariamente ajudaram a realizar o WebMedia 2018.

Salvador, outubro de 2018.

Manoel C.M. Neto (IFBA) – *Coordenador Geral*

Renato Lima Novais (IFBA) – *Vice-coordenador Geral*

Carlos Ferraz (UFPE) – *Coordenador Geral do Comitê de Programa*

Windson Viana (UFC) – *Coordenador Adjunto do Comitê de Programa*

Sumário

Organização do WebMedia 2018	v
Patrocinadores & Apoios	vii
Anais do XVIII Workshop de Teses e Dissertações (WTD)	1
Anais do XV Workshop de Trabalhos de Iniciação Científica (WTIC)	43
Anais do XVII Workshop de Ferramentas e Aplicações (WEA)	85
Anais do V Workshop “O Futuro da Videocolaboração” (WCT-Video) ...	156
Índice de Autores	191

XIV Simpósio Brasileiro de Sistemas Multimídia e Web

**16 a 19 de outubro de 2018
Salvador, Bahia, Brasil**

Coordenação Geral

Manoel C. M. Neto (IFBA) – *Coordenador Geral*
Renato Lima Novais (IFBA) – *Vice-Coordenador Geral*
Carlos Ferraz (UFPE) – *Coordenador do Comitê de Programa*
Windson Viana (UFC) – *Vice-Coordenador do Comitê de Programa*

Coordenador de Publicação

Roberto Willrich (UFSC)

Coordenadores do XV Workshop de Trabalhos de Iniciação Científica (WTIC)

Tiago Maritan (UFPB)
Antônio Maurício Pitangueira (IFBA)
Antônio Carlos Souza (IFBA)

Coordenadores do XVII Workshop de Ferramentas e Aplicações (WFA)

Carlos de Salles (UFMA)
Álan L.V. Guedes (PUC-Rio)

Coordenadores do XVIII Workshop de Teses e Dissertações (WTD)

Eduardo Barrére (UFJF)
Frederico Durão (UFBA)

Coordenadores de Minicursos

Valter Roesler (UFRGS)
Artur Kronbauer (UNIFACS)

Coordenador do V Workshop O Futuro da Videocolaboração

Leandro N. Ciuffo (RNP)
Valter Roesler (UFRGS)
Alexandre Carissimi (UFRGS)

Coordenação da Comissão Especial de Sistemas Multimídia e Web

Adriano C. Machado Pereira (UFMG) – *Coordenador*

Windson Viana (UFC) – *Vice-Coordenador*

Comitê Gestor

Carlos de Salles Soares Neto (UFMA)

Celso Alberto Saibel Santos (UFES)

Fábio de Jesus Lima Gomes (IFPI)

Guido Lemos de Souza Filho (UFPB)

José Valdeni de Lima (UFRGS)

Manoel C.M. Neto (IFBA)

Maria da Graça Campos Pimentel (USP)

Roberto Willrich (UFSC)

Valter Roesler (UFRGS)

Patrocinadores do WebMedia 2018 Patrocinadores & Apoios

Patrocinadores:



Núcleo de Informação
e Coordenação do
Ponto BR



Comitê Gestor da Internet
no Brasil



Em cooperação com:



Organizadores:



Realização:



WebMedia 2018

XXIV Simpósio Brasileiro de Sistemas Multimídia e Web

ANAIS

XVIII Workshop de Teses e Dissertações

Coordenação

Eduardo Barrére (UFJF)

Frederico Durão (UFBA)

XVIII Workshop de Teses e Dissertações (WTD)

Sumário

Prefácio	4
Organização	5
Trabalhos de Doutorado	7
A Hybrid Approach to Recommend Long Tail Items	7
Diogo Vinícius de Sousa Silva (UFBA), Frederico Araújo Durão (UFBA)	
Towards an API for User Attention Prediction in Mobile Notification Overload	13
Alberto Vianna Dias da Silva (UFBA/UFBA), Vaninha Vieira (UFBA)	
Improving the Spatial Keyword Preference Query with Linked Open Data	19
João Paulo Dias de Almeida (UFBA), Frederico Araújo Durão (UFBA)	
Trabalhos de Mestrado	25
Kumo: um serviço para portabilidade em multi-núvens heterogêneas	25
Marcus Rafael Xavier (UFPE), Carlos André G. Ferraz (UFPE), Ioram Schechtman Sette (CESAR)	
A Framework for Automatic Topic Segmentation in Video Lectures	31
Eduardo R. Soares (UFJF), Eduardo Barrére (UFJF)	
Um Modelo de Predição para o Sucesso no Mercado Musical	37
Carlos V.S. Araujo (UFAM), Rafael Giusti (UFAM), Marco A.P. Cristo (UFAM)	

XVIII Workshop de Teses e Dissertações

Prefácio

O Workshop de Teses e Dissertações (WTD) é um dos eventos associados ao WebMedia. Ele é um fórum dedicado à apresentação e discussão de trabalhos de alunos de programas de mestrado e de doutorado, acadêmicos ou profissionais, com foco nas áreas Multimídia e Web. Os trabalhos devem estar em fase intermediária de desenvolvimento.

O objetivo do WTD é propiciar um ambiente dinâmico e construtivo para discussões dos trabalhos, de maneira que os alunos possam apresentar ideias, planejamento e resultados parciais, receber uma avaliação antecipada dos seus trabalhos e obter importantes sugestões para a continuidade e finalização com sucesso de seus mestrados ou doutorados.

O WTD possibilita aos participantes o acesso a um panorama representativo da pesquisa nas áreas Multimídia e Web em diversas instituições e a integração e cooperação entre pesquisadores de diferentes grupos dessas áreas.

Neste ano o evento recebeu um total de 9 submissões de artigos, relacionados a trabalhos de mestrado e doutorado em andamento. Os artigos foram avaliados pelo comitê de programa do evento. Cada artigo recebeu pelo menos dois pareceres, com questionamentos e sugestões que, no caso dos artigos selecionados, foram consideradas para a elaboração das versões finais publicadas nestes anais. Foram três os artigos de mestrado e três os de doutorado escolhidos para apresentação durante o WebMedia 2018 em Salvador-BA. Os artigos serão apresentados e discutidos em sessões específicas, com a participação de membros do comitê de programa e demais pesquisadores participantes.

Agradecemos a todos os autores que submeteram seus trabalhos. Agradecemos aos membros do comitê de programa, pela preciosa colaboração nas revisões dos artigos e nas interações com os apresentadores dos artigos.

Agradecemos ao Professor Manoel C.M. Neto (IFBA), pelo convite para coordenarmos o WTD e pelo seu intenso trabalho e dedicação na organização do WebMedia 2018, e ao Professor Roberto Willrich (UFSC) pela cuidadosa edição desses anais. Agradecemos também à CE-WebMedia da SBC pelo respaldo à nossa indicação feita pelo Prof. Manoel e pela confiança depositada no nosso trabalho.

Salvador, outubro de 2018.

Eduardo Barrére (UFJF)

Frederico Durão (UFBA)

Coordenadores do XVIII Workshop de Teses e Dissertações do WebMedia

Coordenadores do XVIII WTD

Eduardo Barrére (UFJF)
Frederico Durão (UFBA)

Comitê de Programa do XVIII WTD

Carlos André Guimarães Ferraz (UFC)
Cesar Augusto Camillo Teixeira (UFSCAR)
Debora Muchaluat-Saade (UFF)
Frederico Araújo Durão (UFBA)
Marco Antônio Pinheiro de Cristo (UFAM)
Roberto Willrich (UFSC)
Rudinei Goularte (USP)
Valter Roesler (UFRGS)
Vinícius Cardoso Garcia (UFPE)

A Hybrid Approach to Recommend Long Tail Items

Diogo Vinícius de Sousa Silva
Federal University of Bahia
Salvador, Bahia, Brazil
diogovss@ufba.br

Frederico Araújo Durão
Federal University of Bahia
Salvador, Bahia, Brazil
fdurao@ufba.br

ABSTRACT

Techniques in recommendation systems generally focuses on recommending the most important items for a user. The purpose of this work is to generate recommendations focusing on long tail items, and then to conduct the user to less popular items. However, such items are of great relevance to the user. Two techniques from the literature were applied in this study in a hybrid way. The first technique is through markov chains to calculate node similarity of a user item graph. The second technique applies clustering, where items are separated into distinct clusters: popular items (short tail) and non-popular items (long tail). Using the Movielens 100k database, we conducted an experiment to calculate the accuracy, diversity, and popularity of the recommended items. With our hybrid approach we were able to improve the recall by up to 27.97 % when compared to the markov chain-based algorithm, which indicates greater targeting to long tail products. At the same time the recommended items were more diversified and less popular, which indicates greater targeting to long tail products.

KEYWORDS

Recommender System; Long Tail; Graphs; Markov Chain; Clusterization.

1 INTRODUCTION

Most of the methods used by Recommendation Systems (RS) tend to recommend the most popular items to users. Since the majority of users are interested in an item, the likelihood is that a new user will also be interested in that item [8]. By following this logic, it is natural that less popular items are less recommended and consequently less consumed. Usually, the main companies focus their sales on these products envisioning better logistic. If we imagine a company with physical stock, it is easy to understand that it is much cheaper to put the best-selling products on the more evident shelves. However, with the advent of virtual stores the cost of organizing products on shelves is non-existent. In the context of virtual stores has raised the term “infinite-inventory” [1], where products in evidence can be selected according to the preference of each online user. This virtual user will not necessarily have the same preference as other users.

The term “long tail” refers to the set of products not commonly consumed by users [1]. Usually, these products make up the bulk of store stock with low demand. Only a small amount of products contribute for the majority of sales. In contrast, most products (which

would be in long tail) are responsible only for the minority of sales. According to Pareto’s rule [11], 80% of the consequences come from 20% of the causes, that is, 80% of sales would be concentrated in only 20% of the products from the stock. Long tail items are those less popular items, while the other items, i.e. the most popular, will be called “short tail” items.

The long tail phenomenon can also be seen as a way to increase company profits. Generally high popular products are quite marketed by several other companies and of course the competition for sales is great. Since there is a great demand for these products, the price tends to be the lowest due to competition. Therefore, the products profit rate becomes quite low. Considering items with low demand, it is possible to predict a higher profit margin for interested users. The users will be more motivated to pay a higher price due to the low product availability. Another effect of exploring long tail products is the so-called “one-stop shopping convenience” effect. A store that offers long tail products and also popular products delivers an additional convenience for their customers since you find everything you need in one place.

In this context, this work proposes a hybrid technique for improving the recommendation of long tail products. Our approach uses a structure that represents the items and users in a bipartite graph. Then, we determine the proximity through Markov chains using the Hitting Time algorithm. In addition, we also combine the use of this algorithm with a technique of splitting and clustering in order to improve the accuracy of long tail items recommendations. Finally, we indicate the long tail items for recommendations, thus turning the algorithm more assertive and improving the recommendations diversity.

The remainder of the work is organized as follows: Section 2 presents the related works. The Section 3 addresses the proposed long tail recommendation. Section 4 details the experimental evaluation of the proposed technique and discusses the results of the experiments. Finally, Section 5 traces some conclusions and future work.

2 RELATED WORK

Yin et al. [12] has developed four variations of an algorithm for long tail item recommendations. The basic algorithm of the proposal is the Hitting Time, where the users and items are represented in a disjunct, indirect and bipartite graph. From this graph an adjacency matrix is obtained, as shown in Figure 1. The edges of the graph are weighted and represent the relevance of a user’s connection to an item, that is, a user’s rating for that item. To calculate the proximity of unrecorded items, the author calculates the Hitting Time using a type of Markov chains called Random Walk [2]. Using Random Walk, the probability of a user reaching an item not evaluated by him is calculated. The higher the probability, the lower the Hitting Time, and therefore the item should have higher priority in the

recommendation. Transition matrices are derived from a probability matrix. The probability matrix is obtained from the adjacency matrix shown in Figure 1. Shang et al. [10] perform a study of a custom recommendation model using collaborative filtering and ternary relationships, through tripartite graphs, representing users, items, and tags. It proposes a new measure for user similarity based on user tags and preferences. The similarities are calculated using a diffusion-based process and finally compared with recommendations calculated based on similarity of the cosine. Johnson et al. [4] carry out an extension of the work of Yin et al. [12]. Johnson combines the algorithms used by Yin, adapting to tripartite graphs, shown by Shang. Johnson combined the Yin approach with the study of Shang to generate recommendations by collaborative filtering.

Park and Tuzhilin [7] suggest an approach based on splitting and clustering item set. The item set is divided between short tail and various parts of the long tail. Thus, recommendation of short tail items are made based on the individual scores of each item. In the case of the long tail items recommendation, the clusters scores are grouped in each part of the long tail. Park in [6] evolves the previous work by clustering the items based on popularity. The author presents a technique called *adaptive clustering* for recommendations of items according to their popularity. It is possible to define the size and quantities of the clusters in an adaptive way according to the state of the dataset, focusing on items long tail. As a result you get lower error rates and improved performance.

The approach presented in this work uses Hitting Time combined with a technique for item clustering. Our proposal will be based on the Park's approach [7]. We will group the long tail items into clusters according to the average ratings of each item and use this score as weight. The weight represents an additional variable in the Hitting Time algorithm. The main objective is to increase the emphasis on the recommendation of items located in long tail improving the accuracy.

3 PROPOSED SOLUTION

In this section we will explain our hybrid approach composed of Hitting Time algorithm and a clustering technique will turn the recommendations more focused on long tail items.

3.1 Hitting Time

In the Hitting Time algorithm [12] the set of users U and items M are represented in a split-graph and have their corresponding adjacency matrix. Figure 1 illustrates an example of the graph with nodes that could be items or users. The edges of these nodes carry a weight that reflect the user's rating for that item that the edge binds to. When there is no edge connecting two nodes, the weight is 0 (zero). As an example, we have the user node U_1 that has a connection with the item node M_6 . In the matrix of Figure 1 the weight of this edge is 5. In other words, the rating of such a user for item M_6 has a value of 5. The weight of an edge is represented by $a(i, j)$. The variables i and j are the nodes of the graph in an array $A = (a(i, j))_{i, j \in V}$. The variable V represents the set of vertices (nodes) of the graph.

Aiming at calculating the proximity of two nodes in the graph, the algorithm uses a specific type of Markov chain, called random

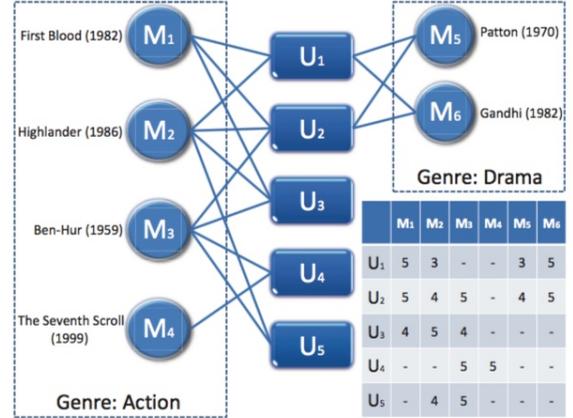


Figure 1: Representation of users and items through a bipartite graph and its respective adjacency matrix [12].

walk. A random walk has a current state (a node in the graph) and with each evolution in time the state is changed, that is, visiting other nodes of the graph. Through the hitting time this path is driven by the weights of the edges and the random walk ends its walk when it arrives at the target node. The algorithm is based on a probability matrix that is computed from the adjacency matrix. Finally, the edge weights are given by Equation $p_{i,j} = P(s(t+1) = j | s(t) = i) = \frac{a(i,j)}{d_i}$, where $d_i = \sum_{j=1}^n a(i,j)$.

Taking into account the probability matrix, a transition matrix is calculated for each chain state until the destination state is reached. Next, the hitting time value may be obtained through of the $H(q|j) = \frac{1}{p_{j,q}} = \frac{\pi_j}{p_{q,j} \pi_q}$, where $\pi_i = \frac{\sum_{j=1}^n a(i,j)}{\sum_{i,j=1}^n a(i,j)}$.

The Hitting Time algorithm is based on the time-reversibility property to guide recommendations with long tail items. This property indicates that the paths are not symmetric. The probability that a node A reaches a node B is different from the probability that node B reaches node A [2]. The algorithm calculation takes the inverse path to calculate the value, that is, from user node to item node. In this way all items, including long tail items, are equally considered for recommendations. Thus, long tail items tend to be more recommended. This fact would not happen if the calculation began in the user-item sense, since there would be more paths that would link to the most popular products.

3.2 Clustering

We divide the set of items between short tail items and long tail items. At this point we use the Pareto's rule [11] as a parameter to separate the most popular items from the less popular ones. All long tail items are clustered considering the score of each item.

Those items with higher scores will have higher priority in the recommendation and will have a greater weight. Such items impact on the decrease in the Hitting Time value. The items with smaller scores are grouped in clusters that will be used to ponder the value of Hitting Time, increasing its size. The value resulting from the

Table 1: Clustering of the item dataset related to the average rating (score) and its respective adjustment factor.

Score Item (S)	Cluster	Adjustment Factor (%)
1 <= S < 2	A	+20
2 <= S < 3	B	+10
3 <= S < 4	C	-10
4 <= S <= 5	D	-20

Table 2: Ordering items based on the Hitting Time algorithm.

Item	Hitting Time	Priority for Recommendation
Titanic	12	1°
Little Dorrit	13	2°
Batman	14	3°
Simple Simon	15	4°
Black	16	5°

application of the adjustment factor to the value found through the algorithm Hitting Time will be called Hitting Time Clustered (HTCL). The value of the adjustment factor that will increase or decrease the Hitting Time is described in Table 1.

To calculate the Hitting Time Clustered (HTCL) we first need to calculate the value to be added. This value may be negative or positive, depending on the cluster in which the item was clustered (see Table 1). This additional value will be added to the previously calculated Hitting Time value. The equation $HTCL(q|j) = H(q|j) + \frac{AF \cdot H(q|j)}{100}$ shows this calculation, where *AF* means the adjusting factor based on Table 1, $H(q|j)$ means the Hitting Time of an item *j* for a user *q* and $HTCL(q|j)$ means the Hitting Time Clustered.

3.2.1 Algorithm In Action. In order to illustrate the algorithm operation, Table 2 presents a list with 5 items recommended for a given user. In our context the items presented are examples of movies. As stated in the previous section, the smaller the hitting time the closer the user is to the item, consequently the item will have higher priority in the recommendation. In Table 2 the items are presented in order of relevance, according to the hitting time.

Now, let’s look at the distribution of items in a chart that represents long tail items and short tail items. Figure 2 represents the point of the set item where splitting is performed and how the items are clustered. In our example, there are 2 items that are presented in the short tail (movies “Titanic” and “Batman”) and 3 items in the long tail (movies “Little Dorrit”, “Simple Simon” and “Black”). For each item in the long tail, its score is shown, i.e. the average of the received ratings. The clustering algorithm will perform similarity classification based on the score (according to Table 1). In this way the movies “Little Dorrit” and “Simple Simon” will be considered similar and will be together in cluster C and the item “Black” will be alone in cluster D. The movies “Titanic” and “Batman” as they are located in the short tail will not be clustered.

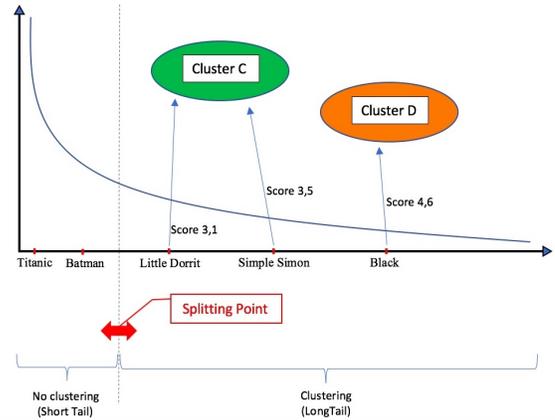


Figure 2: Graphic illustration of splitting and clustering in item dataset.

Table 3: Comparing values between Hitting Time and Hitting Time Clustered.

Item	Hitting Time	Hitting Time Clustered
Little Dorrit	13	11,7
Simple Simon	15	13,5
Black	16	12,8

Table 4: Result after dataset clustering.

Priority for Recommendation	Previous Ranking	Current Ranking
1°	Titanic	Little Dorrit
2°	Little Dorrit	Titanic
3°	Batman	Black
4°	Simple Simon	Simple Simon
5°	Black	Batman

Table 3 lists the 3 items that are located in long tail. For each of them the calculation result is shown by the HTCL approach, based on the adjustment factor shown in Table 1. The order of relevance for the recommendation of the items will no longer be the Hitting Time (HT), but the Hitting Time Clustered (HTCL), instead.

In the Table 4 we see the relevance items order changing after the clustering presented here. We can observe that movie “Titanic” was previously the most relevant recommendation for the user. After the application of clustering we see that movie “Little Dorrit” becomes the most relevant. Notice that in Figure 2 movie “Titanic” is in short tail and movie “Little Dorrit” in longtail. That is, long tail are more prioritized rather than short tail items by just adding the clustering technique. The logic of the Hitting Time algorithm remained the same. We can also observe that the other items also changed position in the ranking of recommendations. Movie “Black” (present in long tail) has increased in the ranking and item “Batman” (present in short tail) fell in the ranking. The item “Simple

Simon” remained in the same position, despite being in the long tail, showing that clustering does not always prioritize the items that are in the long tail.

With such approach the trend is that long tail items with better user ratings take higher priorities. In addition, items that are at the end of the tail can also be recommended more easily and more assertively, since they have a good score.

4 EVALUATION

4.1 Dataset

The dataset used in this study was the MovieLens100k dataset [3]. The MovieLens dataset has approximately 1,682 movies and contains 100,000 ratings on a scale of 1 to 5 scored by around 943 users. The density for the rating matrix is 6.30%, a sparse matrix, which means that most users have not seen most movies. The dataset also includes other information that was not used in this study, such as: age, gender, user occupation, and category of movies.

Each user has rated at least 20 movies. The data was collected through the MovieLens web site (<http://movielens.umn.edu>) during the seven-month period from September 19th, 1997 through April 22nd, 1998. This data has been cleaned up - users who had less than 20 ratings or did not have complete demographic information were removed from this dataset.

4.2 Metrics

4.2.1 Recall. To calculate the accuracy of the proposal, we use a metric called Recall@N, previously adopted by [12]. Recall is the index that indicates the amount of items of interest to the user that appear in the list of recommendations. This index ranges from 0 to 1. The closer to 1 better the recommendation. The metrics idea is to select a user favorite item and include it in a set of randomly selected items. After that, the algorithm ranks all set items and then checks whether the user’s favorite item is displayed in the top@N.

For the recall application we split the dataset into two subset. The first, called test set (or estimation test) was used to randomly select items located in long tail and evaluated with 5 stars. The second subset was used to perform the tests and validation of the generated recommendations. For each item M present in the training subset, we select other items that have not been evaluated by the same user who evaluated item M. Let’s call this training set as B. From there we execute the algorithm and rank all items not evaluated together with the item M. A test case, as described above with item M, is performed several times. In our experiment we varied the number of test cases together with the number of items present in training set B. A number of 500 test cases were executed. Each experiment used a training set B of 100 items.

To calculate the Recall@N we define hit@N like a test case. It is counted how many times the item M appears inside the top@N results, as shown in the following equation $Recall@N = \frac{\sum hit@N}{|L|}$, where |L| is the number of test cases. This way, the higher the recall result, the greater the accuracy of the tested algorithm. Which represents that there are more items of user preference being returned in the top@N results.

4.2.2 Diversity. The metric diversity [5] was used to obtain the degree of distribution of long tail items that are recommended.

With a high diversity the tendency is that long tail items are recommended to the users, that is, the recommendations suffer little influence from very popular items. Such influence causes the discovery of new items hitherto hidden in the long tail.

To calculate diversity we use the same calculation shown in [12]. This metric calculates for a given set of users the top@K items to recommend. Let’s assume that we will use a set of 20 users and for each of them a top@10 ranking, that is, 10 recommended items. In this case we would have a total of 200 recommended items for all users. To calculate the diversity we check how many repeated items are counted only once and then calculate the proportion to the total, as shown in $Diversity = \frac{|\cup I_u \in I|}{|U| \cdot top@N}$, where I_u is the set of single items recommended for all users. The I element means the set of items in the dataset, U represents the set of users and $top@N$ means the number of items recommended for each user.

In our experiments we fixed a set with 200 users and for each user varied the amount of recommended items in top@10, top@20, top@30, top@40 and top@50.

4.2.3 Popularity. The metric popularity represents the quantity of recommended long tail and short tail items. An analysis of this metric together with the others result in a more careful analysis regarding the algorithm performance. This metric calculates the frequency of an item according to the amount of ratings it holds because of the other ratings of dataset [12].

Considering that our dataset has 100,000 ratings, we calculated the popularity of the top@10, top@20, top@30, top@40 and top@50. For each of them we selected 200 different users. The calculation was based on the average popularity of items in each user’s ranking. And for each user the average was calculated to arrive at the final value. That is, we count the amount of evaluations received for each item. And then we relate to the total amount of dataset evaluations, as shown in $Popularity = \frac{R_u}{\sum |R_d|}$, where $R_u = \frac{\sum |R_r|}{|U| \cdot top@N}$, where R_d represents the rating set of the entire dataset. The number of ratings is about 100,000 and the U element represents the set of users selected for the popularity calculation. The $top@N$ element means the amount of recommended items for each user belonging to the U set. In Equation above, R_u represents the rating rate already normalized according to the number of users and recommended items.

The higher the result, the greater the popularity of a particular item. Thus, to have a good result of the algorithm in relation to the recommendation of long tail items, it is interesting that this metric has a low value, that is, recommend little popular items to users. Both the *diversity* metric and the *popularity* metric can measure how far the recommendation is directed to long tail items.

4.3 Baselines

To analyze the effectiveness of our Hitting Time Clustered approach (HTCL), we performed a comparison with 3 different baselines, namely:

- **Hitting Time (HT)** - The algorithm proposed by Yin [12] is presented in Section 3.1.
- **Hitting Time + Clustering All Dataset (HTCA)** - The Hitting Time Algorithm plus clustering similar to our approach. The difference here is in the lack of splitting of

the dataset. That is, there was no separation of items between long tail and short tail and then clustering occurred throughout the dataset.

- **Hitting Time + Clustering Short Tail Dataset (HTCS)** - Also similar to our approach, however, the splitting in the dataset was done in the opposite way. Instead of clustering the items located in the long tail, in this baseline we cluster only the items present in the short tail of the dataset.

4.4 Results

4.4.1 Recall Measurement. Figure 3 shows the performance of our approach (HTCL) compared to the baselines described in the Section 4.3. In the execution presented in the graph were ranked 100 items and then calculated the recall for 500 test cases. We can observe that of all baselines, our approach obtained the best values for the recall in all $top@N$.

Note that when we used only HT, the results were better than any of the other baselines that use clustering, with the exception of HTCL. For example, at the $top@10$ point, the value in baseline HT is 0.086, when we change to our HTCL approach the value was 0.094. In $top@35$ the HTCL outperforms HT in 27,97%. When using HTCA and HTCS, the result was worse than HT in all $top@N$. With the HTCS we had the worst result of all. In $top@5$ HT outperforms in HTCA 31,25% and HTCS in 34,38%. That is, clustering only the short tail will make the recommendations worse than not clustering anything. The best results for the HTCL confirmed the effectiveness of our approach, because the clustering of items in the long tail tend to have a recommendation with better recall.

Other experiments were also performed by varying the number of test cases and the number of ranked items. However the results were similar, following the same order as shown in Figure 3. That is, the best approach was HTCL followed by HT, HTCA and finally HTCS.

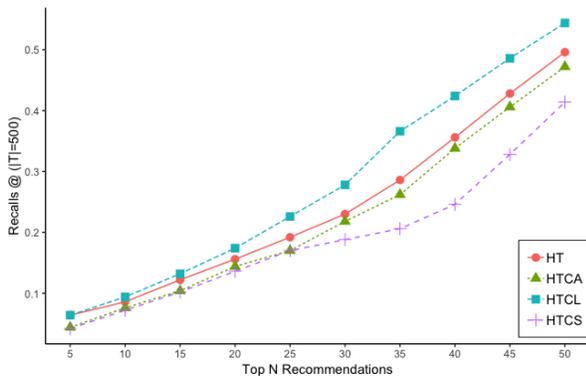


Figure 3: Recall of the top@N items in 500 test cases.

4.4.2 Diversity Measurement. The results of the measurement of diversity were also positive for HTCL. Figure 4 presents the results for an amount of 200 random users when generating top@N recommendations. In this measurement, although the increase is low, it is noticeable mainly when the HTCL recommends from top@30 items. Note that at the top@30, the HTCL approach presents 0.0347,

versus 0.0345 of the HT approach. That is, the items recommended by our proposal have a greater diversity of items. This approach proved to be the best, followed by HT. And following the same order of the recall, the third best was the HTCA and finally the HTCS.

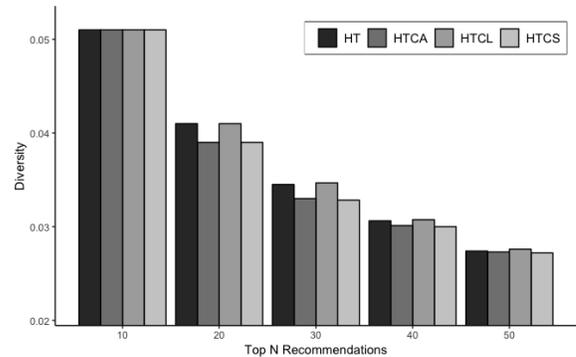


Figure 4: Diversity on Movielens 100k using 200 random users.

4.4.3 Popularity Measurement. In terms of popularity measurements, there was no unanimity. Analyzing Figure 5 we see that the results were quite close to all baselines. The HTCA and HTCS approaches were the ones with the lowest popularity ratings. Our approach got the same result as HT up to the top@20. From then on it managed to overcome HT. See that in top@30 the popularity index in HTCL is 0.3119 and the index in HT is 0.3122. That is, our recommendations tend to suggest less popular products.

Low popularity is not bad for our purpose, since the recommendation of long tail items needs to be more focused on niche products and not among the top sellers. In short, the HTCL approach improves diversity and returns less popular items, which is good, as we are focusing on long tail recommendations. In addition, the good results in diversity and popularity were reached without harming the accuracy, since in the recall method our approach obtained the best results.

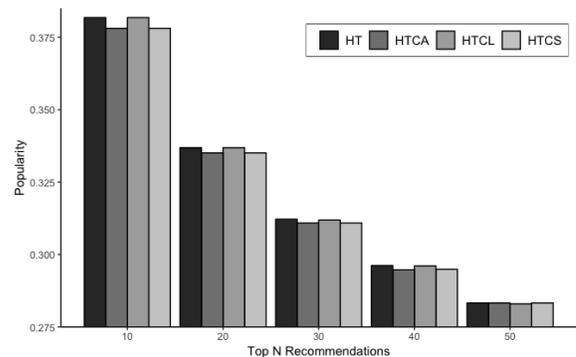


Figure 5: Popularity on Movielens 100k using 500 random users.

4.5 Discussion and Points of Improvements

The approach proposed in this work (HTCL) obtained the best results in the 3 tested metrics. The recommendations tend to be more focused on longtail items, since our approach uses the Hitting Time algorithm together with the clustering technique. The results of baselines HT and HTCL when compared in metrics diversity and popularity show the superiority of our approach. Until reaching the top@20, the two baselines do not show significant differences. The top results of the HTCL approach are evident from top@30, when the HTCL baseline begins to recommend less popular and more diverse items. The results of the other two baselines HTCA (clutter of the entire dataset) and the HTCS (clustering only of short tail) show that clustering if performed in any way can lead to worse results. That is, the clustering performed in these baselines worsened the results already obtained without any type of clustering (baseline HT).

The items diversity was improved and there was a decrease in popularity, but the relevance of the recommendations was not affected. The recall metric allowed us to monitor the accuracy of the recommendations. The HTCL approach obtained the best results at all top@N (see Figure 3). The other baselines that performed the clustering differently, worsened the results of the HT approach. This result has showed the importance of defining a good strategy when clustering a dataset.

Only the variable “item score” was used in the clustering step. Only with this variable the results have already improved, confirming the feasibility of applying different techniques to reach the same goal. Besides the score, there are other variables that can be taken into account in the calculation of similarity, such as: category of the film, producer, and cast. User clustering can also be done through his/her profile data, including age, occupation, gender, among others. These variables are already present in the dataset used in this work and will be the subject of new experiments.

Other techniques can be used in conjunction with a base algorithm and improve recommendations to further explore the long tail. Our approach used Hitting Time as the base algorithm, but other algorithms can be adopted in conjunction with the various clustering techniques. Another possibility is to use, together with clustering, other techniques such as *probabilistic CF algorithm (IRM2)*, *multi-modal similarity and multi-objective evolutionary algorithm (MORS)*, just to name a few.

A limitation is related to the size of the dataset and its sparsity. Datasets with different dimensions and sparsities could generate other results. A deeper analysis could bring new information. The other limitation would be the use of a single domain. We have applied a dataset that aggregates movie ratings, called MovieLens. Analyses in other domains could give different results showing specificities of each tested domain.

5 CONCLUSION

In this work we performed a study to generate recommendations of long tail items. The combination of the Hitting Time algorithm with a clustering technique was adopted in order to give more visibility to long tail items. Using the 100k ratings of MovieLens database, we conducted an experiment to calculate the recall, diversity, and popularity of the recommended items.

The results indicate that two techniques used together may improve the result of a first technique used alone. This procedure should be done carefully, otherwise the effect may be negative. Our proposal presented satisfactory results by focusing on longtail items. We have increased the diversity of recommendations. There was a decrease in popularity of the recommended products. At the same time, the relevance of the recommendations (measured by the recall metric) was also higher compared to the tested approaches. The positive results point out possibilities for retail companies that aim at increasing the profit of their businesses. Since the profit from selling long tail items tend to be larger than short tail items. Focusing part of the sales for these products will bring greater financial returns. This is evidenced by the lower competition for such products. In addition, customers of niche products are usually more loyal and are more willing to pay a higher amount to acquire it.

As future work we intend to extend this work using more variables in the clustering of items, such as age of the user, category of the item, among others. We also intend to adopt other clustering techniques as proposed by [9]. Another possibility is to try out the other algorithms proposed by [12], combining them with the presented clustering technique. Other metrics can be added such as Time, Precision, and F-measure. The use of other databases would also help giving more validity to the results, as well making comparisons with other baselines.

REFERENCES

- [1] Chris Anderson. 2006. *The Long Tail: Why the Future of Business Is Selling Less of More*. Hyperion.
- [2] Gunter Bolch, Stefan Greiner, Hermann de Meer, and Kishor S. Trivedi. 1998. *Queueing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications*. Wiley-Interscience, New York, NY, USA.
- [3] F. Maxwell Harper and Joseph A. Konstan. 2015. The MovieLens Datasets: History and Context. *ACM Trans. Interact. Intell. Syst.* 5, 4, Article 19 (Dec. 2015), 19 pages. <https://doi.org/10.1145/2827872>
- [4] Joseph Johnson and Yiu-Kai Ng. 2017. Enhancing Long Tail Item Recommendations Using Tripartite Graphs and Markov Process. In *Proceedings of the International Conference on Web Intelligence (WI '17)*. ACM, New York, NY, USA, 761–768. <https://doi.org/10.1145/3106426.3106439>
- [5] Matev Kunaver and Toma Porl. 2017. Diversity in Recommender Systems A Survey. *Know-Base Syst.* 123, C (May 2017), 154–162. <https://doi.org/10.1016/j.knosys.2017.02.009>
- [6] Y. J. Park. 2013. The Adaptive Clustering Method for the Long Tail Problem of Recommender Systems. *IEEE Transactions on Knowledge and Data Engineering* 25, 8 (Aug 2013), 1904–1915. <https://doi.org/10.1109/TKDE.2012.119>
- [7] Yoon-Joo Park and Alexander Tuzhilin. 2008. The Long Tail of Recommender Systems and How to Leverage It. In *Proceedings of the 2008 ACM Conference on Recommender Systems (RecSys '08)*. ACM, New York, NY, USA, 11–18. <https://doi.org/10.1145/1454008.1454012>
- [8] Francesco Ricci, Lior Rokach, and Bracha Shapira. 2011. Introduction to Recommender Systems Handbook. In *Recommender Systems Handbook*. Springer, 1–35.
- [9] Jia Rongfei, Jin Maozhong, and Liu Chao. 2010. A new clustering method for collaborative filtering. In *2010 International Conference on Networking and Information Technology*. 488–492. <https://doi.org/10.1109/ICNIT.2010.5508465>
- [10] Ming-Sheng Shang, Zi-Ke Zhang, Tao Zhou, and Yi-Cheng Zhang. 2010. Collaborative filtering with diffusion-based similarity on tripartite graphs. *Physica A: Statistical Mechanics and its Applications* 389, 6 (2010), 1259 – 1264. <https://doi.org/10.1016/j.physa.2009.11.041>
- [11] Kazuhiro Yamashita, Shane McIntosh, Yasutaka Kamei, Ahmed E. Hassan, and Naoyasu Ubayashi. 2015. Revisiting the Applicability of the Pareto Principle to Core Development Teams in Open Source Software Projects. In *Proceedings of the 14th International Workshop on Principles of Software Evolution (IWPE 2015)*. ACM, New York, NY, USA, 46–55. <https://doi.org/10.1145/2804360.2804366>
- [12] Hongzhi Yin, Bin Cui, Jing Li, Junjie Yao, and Chen Chen. 2012. Challenging the Long Tail Recommendation. *Proc. VLDB Endow.* 5, 9 (May 2012), 896–907. <https://doi.org/10.14778/2311906.2311916>

Towards an API for User Attention Prediction in Mobile Notification Overload

Alberto Vianna Dias da Silva
Federal Institute of Bahia (IFBA)
Jacobina, Bahia, Brazil
Graduate Program in Computer Science (PGCOMP),
Federal University of Bahia (UFBA)
Salvador, Bahia, Brazil
profalbertovianna@gmail.com

Vaninha Vieira
Graduate Program in Computer Science (PGCOMP),
Department of Computer Science, Federal University of
Bahia (UFBA)
Salvador, Bahia, Brazil
vaninha@ufba.br

ABSTRACT

Mobile device users receive a lot of notifications per day. Some important notifications may be lost because of notification overloading. For some user, losing a notification can generate serious problems. This work proposes an Application Programming Interface (API) that uses context-awareness and machine learning to predict user attention to support mobile application developers and increases the chances of a important notification to be attended.

KEYWORDS

Mobile notification, Mobile sensing, Context-awareness, User attention prediction, Attention awareness

1 INTRODUCTION

Mobile devices are now part of everyone's life. The enormous number of mobile subscribers can give an idea of that. From 1989 to 2015 this number increased from seven million to 7.1 billion of mobile subscribers [13]. In addition, most of these subscribers use a smart mobile device, such as: smartphone, tablet, smartwatch, etc.

The possibility of notifying a user and get his/her attention is an interesting feature that smart mobile devices have. Each application inside these devices can send a mobile notification to users and give an information to them.

A notification can be defined as a visual, audible, or tactile signal, generated by an application or service, that gives information to a user who is out of focus [6]. On mobile devices, notifications are usually sent instantly when there is some activity that may be relevant to the user when the application is not open, such as: a new email, a text message that has just arrived, or a new comment on their social networks. In some cases the user takes immediate action after the notification arrives, while in others it is simply ignored. These actions depend on the relevance of the notification and user context [10].

On mobile devices a notification is a message that can be displayed outside the application context. When the application issues a notification, it first appears as an icon in the notification area.

Mobile notifications have been used for different purposes. Some examples are: to tell the user that a new email has arrived, to alert

the user a new instant message needs to be responded, or even to notify that the device is fully charged.

In a study published by Falaki *et al.* (2010) [4], they measure the average of applications installed in mobile devices from 255 users. This number varies from 10 to 90 applications and average was roughly 50. Hence, an average of 50 applications can notify the user. Yahoo Aviate (2014) [17] collected data from their application use and disclosed more recent information about Android users. They informed that users have an average of 95 applications installed and an average of 35 applications that are used through the day.

Iqbal and Horvitz (2010) [6] evaluated the use of notifications generated by an email tool in a desktop computer. The study mentioned that users found notifications very important for them, however inopportune and disruptive in some moments - imagine users receiving notification from 50 different applications.

Because of that, the number of notifications a user receives in mobile devices has increased a lot and some problems have emerged, such as: receiving notifications in inopportune moments; interrupting an important task at work; not realizing an urgent e-mail has arrived and not responding because of the big amount of notifications received before, etc. Thus, resulting in a scenario very different from Weiser and Brown (1997) [15] calm computing technologies.

One of the most significant scenarios is the notification considered important by the user. For example, a businessman who owns a company which sells products over the Internet, considers a notification that a product could not be delivered a very important one. Also, for an elderly, a notification warning it is time to take some medicine is very important. In current scenario, there is a big chance that a user does not see that notification because of the overload of notifications, and mobile application developers have limited resources to deal with this situation.

A usual way to handle with lots of notifications is to group them with some linkage process, as used by [7]. Our study uses a different approach. We used context-awareness and machine learning to predict user attention.

We believe that context information about the user could infer the user attention to incoming notifications. For instance, a physical activity can give suggestions if a user is paying attention to her/his phone.

Context is defined by Dey [3] as "any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves"

In: XVIII Workshop de Teses de Dissertações (WTD 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pôsteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.
© 2018 SBC – Sociedade Brasileira de Computação.
ISBN 978-85-7669-435-9.

Context information is important to define the user current state and the environment where it is inserted, but only this is insufficient. In order to use this information satisfactorily, it is ideal to build an adaptive system that meets the needs of the user in real time using context information. In short, a context-aware system.

Context-aware systems are able to adapt their operations to the current context, without explicit intervention of the user, aiming to increase their usability and effectiveness by taking into account context elements [1]. A system is context-aware if "it uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task" [3].

A large-scale study was published by Shirazi *et al.* (2014) [10]. One of their proposed future work was the automatic assessment of notifications, to deliver notification when user appreciated and had the least negative impact. This can be achieved by using context information about user, network, and environment.

One motivation of this work has a real scenario. We developed a solution for monitoring beta-blocked cardiac patients [11], when we evaluated the application, some important notifications were lost by the user because of the overloading of notifications. The solution needed an intelligent mechanism to raise the chances for user seeing a notification and lead to the characterization of the research problem presented in next section.

1.1 Related Works and Research Problem

Recent studies in the field have already tried to improve notification process [2, 8, 9]. However, they used offline classifying algorithms from machine learning to try to improve notification process in predicting user attention.

We believe that offline classifying algorithms are not the best choice for this task, because notification context is very dynamic. For example, if user installs a new application, the learning model could give a wrong result, because it will know nothing about the new application and user habits. Thus, an online machine learning technique is more suitable for such task. Corroborating with this idea, Mehrotra *et al.* (2015) [8] attempted to use online learning. However, they still used offline learning, because the learning model would be re-trained once a day without further analysis.

Furthermore, these studies mentioned before conceived context models, analyzed them, but did not improve them in any way. Therefore, there is no context model as a pattern to be used in attention prediction to improve notification process.

In addition, none of the studies created any software component to facilitate the use of the proposed mechanisms. Developers, who want an intelligent mechanism of notification, could not use existing solutions. They need an Application Programming Interface (API), a service, a framework, or a library to incorporate those mechanisms in their applications.

The term API, in this work, is the same used in Stylos and Myers (2007) [12], can refer to either the abstract notion of the interfaces, the distribution suitable for use when programming (binaries with definition files), and the implementation source of the API. We choose to develop an API because we do not want to change the way of writing a program, as a framework do, and our proposal is collection of software components that differs our solution from a library.

The research gaps showed before lead to the research question presented in Table 1.

Table 1: Research Question.

How to improve the probability of a notification generated by a mobile application to be attended by the user?

Our study aims to present a solution that uses a mature context model and online machine learning to predict user attentiveness. We used the same concept for the word "attentiveness" as used by Pielot *et al.* (2014) [9] extended to notifications: "we define user attentiveness as the degree to which (s)he is paying attention to incoming instant messages".

1.2 Objectives

The research question that conducts this work is presented in Table 1. From the research question, the general objective emerged:

- Create an API to add intelligence in mobile notification process to predict user attentiveness, using context-awareness and online machine learning, to support mobile developers and increase the chances of a notification be attended.

To reach the general objective, the next steps were performed as specific objectives:

- Investigate the state-of-art of the area and conceive a discussion about topic;
- Create a context model with the contextual elements for the basis of the proposed intelligent mechanism;
- Investigate which machine learning algorithm is most suitable to the scenario;
- Validate whether the context model can be used to predict user attention using selected machine learning algorithms in an experimentation study;
- Design and develop a novel API to predict attention in mobile notifications using online machine learning;
- Perform an experiment to validate the developed solution.

2 METHODOLOGY

This section describes the methodology proposed for this work. A complete overview of it can be achieved in Figure 1.

The first task performed to reach the objectives of this research was to conduct a experimentation study to check the viability of the proposal. After the viability validation and promising results found, we went further on the definition of the methodology.

There are different artifacts proposed in this study. One of the main artifacts is a novel context model to be used as a reference in the area. To reach this goal a step called "Defined Context Model" is presented in Figure 1. The context model will be defined through literature review and related works analysis, during viability experimentation a first version of the model was defined (more information on Section 4). Other versions will be produced through the use of focus group, participants interviews, specialists surveys, and a systematic mapping study.

This model will be used to gather real data from different mobile devices users to construct a real data set. Machine learning algorithms will be applied to check which algorithm has the best

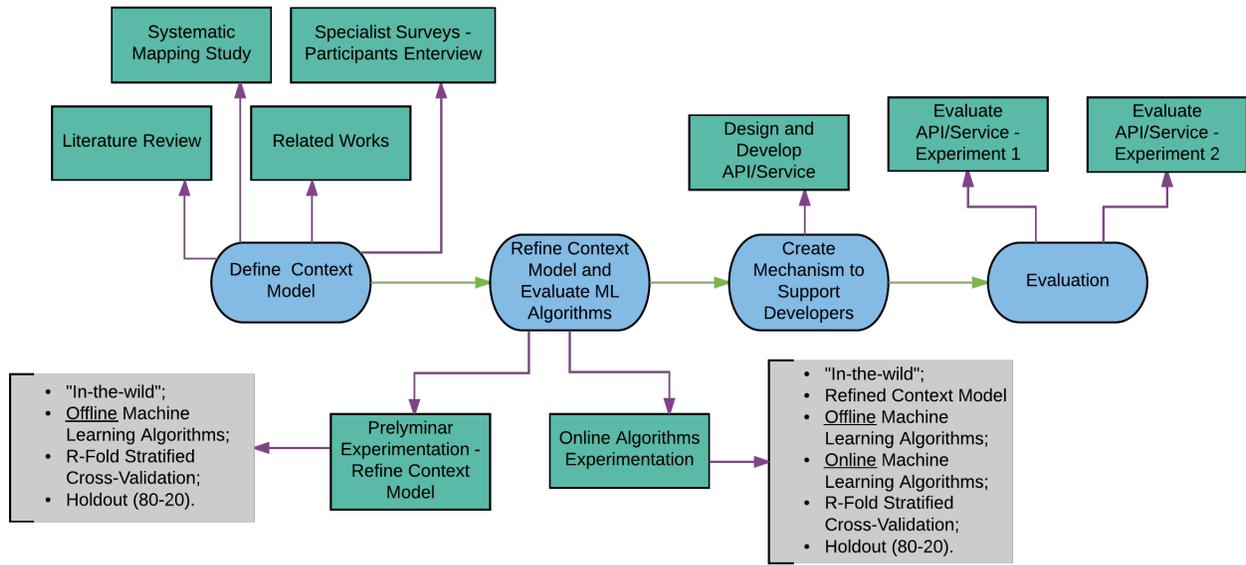


Figure 1: Methodology Overview.

results for this work scenario. This step is represented in Figure 1 as "Refine Context Model and Evaluate ML Algorithms". Two experimentations are expected, both will use as machine learning validation methods r-fold stratified cross-validation and holdout.

The step called "Create Mechanism to Support Developers" represents the development of the API/service using the context representation, the machine learning algorithm with the best result, and the architecture presented in Figure 2. **The main idea of the solution is to respond: will the user see the notification or not?**

The solution will be evaluated in two controlled experiments. At first, a preliminary experiment inside the academy. The second, a controlled experiment, with the evolutions and the resolution of unforeseen problems of the first experiment. In both experiments, two application that generates important notifications will be used as study cases and we will check whether the solution could predict user attentiveness.

3 CURRENT STATUS OF WORK

At the current moment, we are evolving the context model. We applied a focus group with 15 mobile device user to extract new context elements. Furthermore, we are conducting a systematic mapping to identify and characterize the body of knowledge of scientific literature reporting solutions that support the use of context-awareness in notification process in mobile and wearable devices.

The main research question of the study is "How has context-awareness been used to support notification overloading in mobile and wearable devices?" For now, we have already surpassed the selection/filtering phase and we are extracting the selected paper presented in Table 2.

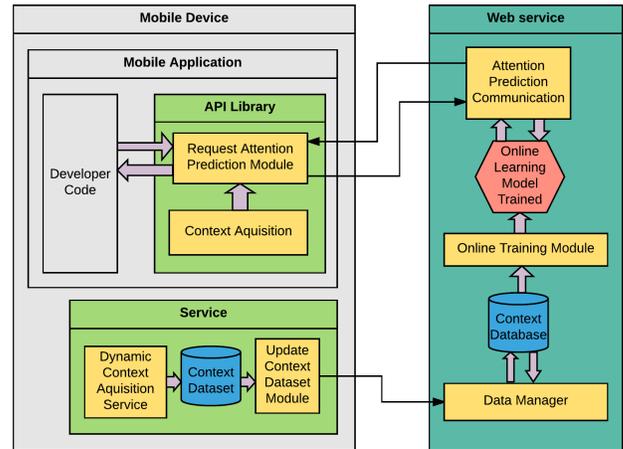


Figure 2: The proposed Architecture.

Table 2: Mapping study selection results.

Results	Amount of paper
Duplicated papers	445
Rejected papers	623
Accepted papers	38

4 PREVIOUS RESULTS

A preliminary study was performed to validate the viability of the project. A context model was constructed after collection of all

different contextual elements from the studies that had a relevant impact on results. A few more contextual elements we considered important were also added and grouped in the first version of a new context model. The context model has contextual elements that can be related to prediction of user attention.

4.1 Context Model - First Version

In order to model a context-aware application, it is necessary to identify entities and contextual elements [14]. The context model used in this preliminary study is presented in Table 3. A contextual element can be static (does not change over time), dynamic (changes over time), explicit (informed by the user) and implicit (obtained without user interaction).

Table 3: First version of a High-level Context Model.

Entity	Contextual Element		Data Source
	Name	Data Type	
User	Location	Coordinate	GPS / WIFI
User	Physical Activity	Enum.	Google API
User	Last Time Near	Time	App, sensor and Clock
Smartphone	Last Time Unlocked	Time	App and Clock
Smartphone	Connection Type	Enum.	Smartphone Info
Smartphone	Ringer Mode	Enum.	Smartphone Info
Environment	Noise Level	Number	Microphone
Environment	Day of Week	Enum.	Calendar
Environment	Day-Shift	Enum.	Clock
Notification	Application name	Text	Smartphone Info
Notification	Title of notification	Text	Smartphone Info
Notification	Attended	Boolean	Smartphone Info

In context model presented in Table 3, four entities were identified: the user that received the notification; the smartphone that generated the notification; the environment the user was immerse; and data from the notification itself. All contextual elements were classified as dynamic and implicit.

The smartphone is the unique data source used in the project. In this way, Table 3 shows which element of smartphone was used to acquire context information.

After definition of context model, we performed a preliminary experimentation study and developed a smartphone application to gather information and contextual elements from different users. We named this application “DataCollector”.

4.2 Preliminary Experimentation Study

DataCollector was developed in native Android SDK¹. We implemented a NotificationListenerService² to monitor user notifications in an unobtrusively way. In favor of identifying the physical activity of the user, we used Activity Recognition API for Android from Google³.

When a notification is presented, following information is collected: location latitude and longitude from global positioning system (GPS) sensor; surrounding sound amplitude using microphone; and last time user was close to the phone in seconds, we gather this information from proximity sensor.

Other information acquired are: last time user was interacting with her/his device, in seconds, determined by last screen lock moment; connection type (mobile, wifi); ringer mode (vibrating, ringing, etc.); day-shift and day of week were extracted from notification posted moment.

A Notification has several attributes, we used in this study the notification title. Notification message and notification ticker were not used due to privacy concerns.

To the best of our knowledge, there is no way to get directly the information that a user attend to a notification or not. To infer the user attended the notification, we implemented two features. First, we check for 30 seconds whether the application that generated the notification had opened after the notification was removed. If application opened, we consider the user clicked the notification. Second, clicking a notification is not the only way to respond a notification, user can unlock the smartphone device and attend or respond the notification without clicking the notification. Considering this scenario, we use the same strategy performed in literature [8, 9], if the user opens the application that generated the notification in 10 minutes, since after it arrival, notification is set to “attended”.

4.2.1 Preliminary Predicting Attentiveness Model. We conducted a data collection and installed DataCollector application in 3 people different devices. They used the application for at least 15 days. All of them are Brazilians and volunteered for the study. DataCollector collects all kinds of notification: instant messages, e-mails, system notifications, etc. At the end, 33892 notification were collected. After pre-processing, 10584 notification were analyzed.

4.2.2 Evaluation. The evaluation was conducted using both methods R-Fold Stratified Cross-Validation and Holdout. For the first one, we use R=10 proportional folds to validate. The second one, we use the proportion of 80% as training data-set and 20% as test data-set. We perform the experiment using Weka machine learn software [5]. Tested machine learning algorithms were: (i) C4.5; (ii) Random Forest; (iii) AddaBoost; (iv) Naive Bayes; (v) KNN (K equal to 1, 2, and 3); (vi) SVM; and (vii) Voted Perceptron. Each algorithm and validation method combination was configured to test with the execution of 100 interactions. Table 4 shows results with correct classification rate average for each algorithm and validation method. Random forest had the best results in R-Fold Stratified Cross-Validation and Holdout.

¹<https://goo.gl/J5wb1X>

²<https://goo.gl/Oh6B5o>

³<https://goo.gl/214fMQ>

Table 4: Algorithms average of correct classification rates.

Algorithm	R-Fold	Holdout
Random Forest	93.87%	93.83%
KNN (K=1)	93.84%	93.82%
KNN (K=2)	93.36%	93.24%
C4.5	93.01%	92.81%
KNN (K=3)	92.89%	92.68%
SVM	79.93%	80.02%
Voted Perceptron	77.94%	77.85%
Naive Bayes	74.43%	74.50%
AdaBoost	71.05%	71.27%

Some algorithms, such as KNN, SVM and Perceptrons, use attributes with numeric values in their logic. As Weka stores all attributes values as double floating point numbers internally [16], we decided to not transform all categorical attributes in numeric, allowing Weka transforms by its native implementation.

5 EXPECTED CONTRIBUTIONS

As expected contributions, we can mention:

- A novel and mature context model for notification attentiveness;
- An API that can be used to help mobile developers in attention prediction of the user;
- A smartphone application for collect raw data about notifications;
- A raw data-set of notifications that can be used in future studies.

ACKNOWLEDGMENTS

The authors would like to thank all the volunteers who participated in the experiment. The work is partially supported by the CAPES Agency under Grant No.:#1490744.

REFERENCES

- [1] Matthias Baldauf, Schahram Dustdar, and Florian Rosenberg. 2007. A survey on context-aware systems. *International Journal of Ad Hoc and Ubiquitous Computing* 2, 4 (2007), 263.
- [2] Fulvio Corno, Luigi De Russis, and Teodoro Montanaro. 2015. A context and user aware smart notification system. In *IEEE 2nd World Forum on Internet of Things (WF-IoT)*. Milan, Italy, 645–651.
- [3] Anind K. Dey. 2001. Understanding and using context. *Personal and ubiquitous computing* 5, 1 (2001), 4–7.
- [4] Hossein Falaki, Ratul Mahajan, Srikanth Kandula, Dimitrios Lymberopoulos, Ramesh Govindan, and Deborah Estrin. 2010. Diversity in Smartphone Usage. In *MobiSys '10 Proceedings of the 8th international conference on Mobile systems, applications, and services*. San Francisco, 179–194.
- [5] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H Witten. 2009. The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter* 11, 1 (2009), 10–18.
- [6] Shamsi T Iqbal and Eric Horvitz. 2010. Notifications and Awareness : A Field Study of Alert Usage and Preferences. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work, CSCW*. New York, New York, USA, 27–30.
- [7] Juan Antonio Lossio-Ventura, Hakim Hacid, Mathieu Roche, and Pascal Poncelet. 2016. Communication overload management through social interactions clustering. In *Proceedings of the 31st Annual ACM Symposium on Applied Computing*. Pisa, Italy, 1166–1169.
- [8] Abhinav Mehrotra, Mirco Musolesi, Robert Hendley, and Veljko Pejovic. 2015. Designing Content-driven Intelligent Notification Mechanisms for Mobile Applications. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. Osaka, Japan, 813–824.
- [9] Martin Pielot, Rodrigo De Oliveira, Haewoon Kwak, and Nuria Oliver. 2014. Didn't You See My Message? Predicting Attentiveness to Mobile Instant Messages. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Toronto, Ontario, Canada, 3319–3328.
- [10] Alireza Sahami Shirazi, Niels Henze, Tilman Dingler, Martin Pielot, Dominik Weber, and Albrecht Schmidt. 2014. Large-Scale Assessment of Mobile Notifications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Toronto, Ontario, Canada, 3055–3064.
- [11] Alberto Vianna Dias da Silva and Vaninha Vieira. 2016. BBAware - A Context-Aware Mobile and Wearable Architecture for Monitoring Beta-Blocked Cardiac Patients. In *Simpósio Brasileiro de Computação Ubíqua e Pervasiva*. Anais do XXXVI Congresso da Sociedade Brasileira de Computação, Porto Alegre.
- [12] Jeffrey Stylos and Brad Myers. 2007. Mapping the space of API design decisions. In *Proceedings - IEEE Symposium on Visual Languages and Human-Centric Computing, VL/HCC 2007*. Coeur d'Alene, Idaho, USA, 50–57.
- [13] TeleGeography. 2015. Global Mobile Subscribers Surpass 7 Billion. <https://www.telegeography.com/press/press-releases/2015/07/28/global-mobile-subscribers-surpass-7-billion/index.html>. (2015). Online; accessed 08 February 2017.
- [14] Vaninha Vieira, Patricia Tedesco, and Ana Carolina Salgado. 2011. Designing context-sensitive systems: An integrated approach. *Expert Systems with Applications* 38, 2 (2011), 1119–1138.
- [15] Mark Weiser and John Seely Brown. 1997. The Coming Age of Calm Technology. In *Beyond Calculation*. Springer New York, New York, 75–85.
- [16] Weka. Why am I missing certain nominal or string values from sparse instances? go.gl/y9DmO. (????). Online; accessed 24 Sempember 2016.
- [17] YahooAviate. 2014. How Android Users Interact with Their Phones. <https://yahooviate.tumblr.com/post/95795838933>. (2014). Online; accessed 30 March 2017.

Improving the Spatial Keyword Preference Query with Linked Open Data

João Paulo Dias de Almeida
Federal University of Bahia
Av. Adhemar de Barros, Ondina
Salvador, Bahia 40170-115
joao.dias@ufba.br

Frederico Araújo Durão
Federal University of Bahia
Av. Adhemar de Barros, Ondina
Salvador, Bahia 40170-115
fdurao@ufba.br

ABSTRACT

This paper presents a Spatial Keyword Preference Query (SKPQ) enhanced by Linked Open Data. This query selects objects based on the textual description of features in their neighborhood. The spatial relationship between objects and features is explored by the SKPQ using a Spatial Inverted Index. In our approach, the spatial relationship is explored using SPARQL. However, the main benefit of using SPARQL is obtained by measuring the textual relevance between features' description and user's keywords. The object description in Linked Open Data is much richer than traditional spatial databases, which leads to a more precise similarity measure than the one employed in the traditional SKPQ. We present an enhanced SKPQ and two experimental evaluations of the proposed approach, comparing it with the traditional SKPQ. The first conducted experiment indicate a relative NDCG improvement of the proposed approach over the traditional SKPQ of 20% when using random query keywords. The second experiment shows that using real query keywords, our approach obtained a significant increase in the MAP score.

KEYWORDS

Spatial data, Query evaluation, Query processing, Linked Open Data

1 INTRODUCTION

Many preference queries specify the user preference using query keywords. For instance, a user looking for a Japanese restaurant can specify his preference with the query keywords "japanese restaurant". These queries evaluate an object as relevant for the user when it's textual description shares words in common with the query keywords [3, 4]. In this way, the more words in common, the better the textual relevance between an object and the query keywords. However, this evaluation method has limitations, especially to objects with short textual descriptions. It is possible that an object is relevant for the user but its description does not contain any query keyword. One possible solution for this problem is offering a wider textual description for the objects.

Motivated by this problem, we use the data available at Linked Open Data (LOD) cloud to enrich the textual description of objects.

A large number of researches have recently studied how to improve the object's textual description using the LOD cloud. This improvement is applied in several areas of research, such as Recommender Systems [6, 9] and Information Retrieval [1, 11]. However, to the best of our knowledge, we are the first to apply a similar improvement in a Spatial Keyword Preference query.

This paper proposes a location-based solution that exploits the benefits of a LOD dataset for enriching the object textual description. We employ our solution at Top-k Spatial Keyword Preference Query (SKPQ) [5]. This query accesses objects from a traditional database like OpenStreetMap. However, a LOD database like DBpedia contains objects' descriptions wider than the ones available at OpenStreetMap. The contributions of this work are a novel semantic model for enhancing the SKPQ and an analysis on how the wider textual description influences the query results.

The remainder of this paper is structured as follows: Section 2 introduces a motivating scenario; Section 3 describes the SPARQL queries employed, then Sections 4 and 5 present the experimental evaluations and the discussion about these evaluations. Finally, Section 6 presents the conclusions and future work.

2 MOTIVATING SCENARIO

The Top-k Spatial Keyword Preference Query (SKPQ) is a preference query that uses query keywords to describe the user preference [5]. The SKPQ searches for spatial objects of user's interest based on spatio-textual objects¹ of reference (features) in their spatial neighborhood. For example, Figure 1 describes a spatial area with spatial objects p (e.g. hotels) and features f (e.g. any establishment). Consider a user interested in book a hotel close to a Japanese restaurant. The user specifies the query keywords "japanese restaurant" and the spatial selection criteria (represented by the circle around the objects p). An evaluation method defines that the textual description of the object f_1 "restaurant" has textual relevance to query keywords. However, the textual description of object f_4 "japanese restaurant" is more textual relevant because it has the same words as the query keywords. Objects f_2, f_3, f_5, f_6, f_7 have no textual relevance to the query keyword, while f_5 does not satisfy the spatial selection criteria too. The SKPQ returns the object p_3 as the best hotel for the user's need, since f_4 has the greatest textual relevance among all features and satisfies the spatial selection criteria.

Suppose a SKPQ with query keywords "oriental food". Considering Figure 1, this query does not return any objects. Neither

In: XVIII Workshop de Teses de Dissertações (WTD 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pósteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.
© 2018 SBC – Sociedade Brasileira de Computação.
ISBN 978-85-7669-435-9.

¹Spatio-textual object is an object with spatial coordinates (e.g. latitude and longitude) and text.

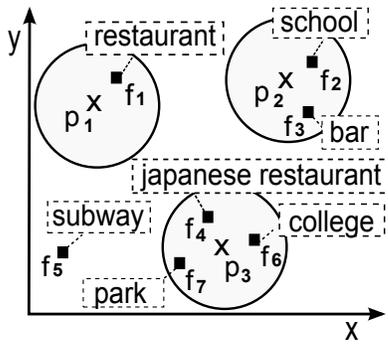


Figure 1: Spatial objects of interest (p) and features (f) associated with their textual descriptions.

the word “oriental” or “food” are present in any textual description. Note that “oriental food” has semantic relevance to “japanese restaurant”, but the evaluation method is not able to identify this relationship. In this example, the query fails to retrieve relevant objects when query keywords are “oriental food”. So, we propose a solution using a LOD dataset to enhance the object textual description, in order to achieve better object evaluation. A wider textual description for objects f can improve the object evaluation. If object f_4 had a better textual description, the word “food” or “oriental” might appear in the textual description. In this scenario, the semantic relationship offered by the LOD dataset can be very helpful too.

3 SPARQL

SPARQL is a query language that can be used to express queries across diverse data sources. The data queried using SPARQL might be stored natively as RDF or viewed as RDF via middleware. A SPARQL endpoint is used to enable users to query a knowledge base via the SPARQL query language. DBpedia and LinkedGeoData endpoints can be accessed at <http://dbpedia.org/snorql/> and <http://linkedgeo.org/sparql>. Listing 1 introduces a SPARQL query to obtain features within 200 m from an object of interest. In Listing 1, *objectURI* is a URI to an object of interest.

The predicate *geo:geometry* is defined at Geo-SPARQL [12], an ontology that represents features and geometries. In Listing 1, the variable *location* matches with the spatial coordinates of objects around an object of interest. The function *bif:st_intersects()* returns true if there is at least one point in common between the spatial coordinates *location* and *sourcegeo*. The tolerance for the matching in units of linear distance is supplied at the third parameter of *bif:st_intersects()*. The tolerance is 200 m as illustrated at Listing 1.

4 EXPERIMENTAL EVALUATION

In this section, we present our methodologies and the results obtained during the experimental evaluation. In addition, we discuss the dataset and the methodologies employed to analyze the proposed approach. The experiments were performed in two ways, each with a unique methodology. In the first experiment, the users’ ratings from Google Maps were extracted to evaluate the queries

```
SELECT DISTINCT ?resource WHERE {
  ?objectURI geo:geometry ?sourcegeo.
  ?resource geo:geometry ?location ;
  rdfs:label ?label .
FILTER( bif:st_intersects( ?location, ?sourcegeo,
0.2 ) ) . }
```

Listing 1: SPARQL query to find features that satisfies the spatial selection criteria.

```
SELECT * WHERE {
  ?var rdfs:label "OSMlabel" .
  ?var geo:lat ?lat.
  ?var geo:long ?lon. }
```

Listing 2: SPARQL query to obtain the objects of interest to process SKPQ-LD.

result. In the second experiment, the users’ ratings were extracted from TripAdvisor².

4.1 Datasets

In this work, we used three datasets to process the SKPQ. The OpenStreetMap (<http://www.osm.org>) dataset was used to process SKPQ and, DBpedia and LinkedGeoData were used to process SKPQ-LD. Additionally, two publicly available datasets were used to evaluate the obtained query results: the Google Maps dataset and the OpinRank dataset.

Extracts are pieces of OpenStreetMap data pruned at the region of individual continents, countries, or metropolitan areas. Mapzen³ maintains updated extracts for many cities. In this work, we used Mapzen to obtain OpenStreetMap data from Dubai. We process this dataset to extract only spatio-textual objects. The set of objects of interest P is composed by spatial objects whose the category in the OpenStreetMap is hotel, while the set of features F is composed by the other spatio-textual objects. The OpenStreetMap extract representing Dubai generated 162 objects of interest, 2243 features, 1906 unique terms and 12256 terms in total.

LinkedGeoData uses the information collected by the OpenStreetMap project and makes it available as an RDF knowledge base according to the Linked Data principles. To process SKPQ-LD we used SPARQL at LinkedGeoData to obtain a set of objects P equivalent to the one obtained from Mapzen, as illustrated by Listing 2. This SPARQL query returns a list of objects with the same name as the one stored at Mapzen, but different spatial coordinates (i.e. there are several places called “McDonald’s” in Dubai, but at different spatial coordinates). Then, we selected only the object with the same name and the same spatial coordinate as the one selected as p object at Mapzen. Additionally, we used the LinkedGeoData endpoint to access feature’s textual description. The textual description obtained from LinkedGeoData is composed by *rdf:type* and *rdfs:label* predicates.

²<https://www.tripadvisor.com.br/>

³<https://mapzen.com/data/metro-extracts/>

Hotel name	Aspect Rating Value
Hatta Fort Hotel	4.107
Al Manzil Hotel	4.341
Park Hyatt	4.342

Table 1: Example of information available in OpinRank dataset related to the query “great location”.

Both DBpedia and LinkedGeoData have public access. We accessed the data from their respective endpoints, storing the obtained data in a local repository. When the query searches for the textual description of one object, it first searches in the local repository. If the search fails, it looks for the information in the endpoints.

4.1.1 Dataset for Experiment 1. Besides the datasets used to process the SKPQ and SKPQ-LD, we used the Google Maps dataset and OpinRank dataset to evaluate the queries. The Google Maps dataset was accessed through the Google Places API. This dataset contains objects of interest that are updated frequently through owner-verified listings and user-moderated contributions. We extract from Google Maps the users’ ratings to the hotels retrieved by the SKPQ and SKPQ-LD. These users’ ratings are used to evaluate both SKPQ and SKPQ-LD.

4.1.2 Dataset for Experiment 2. The OpinRank dataset [7] contains hotel reviews and aspect ratings. There are 5 aspects ratings related to hotels: *cleanliness*, *value*, *service*, *location* and *room*. The aspect ratings values are on a scale of 1-5. Ganesan and Zhai [7] manually created textual queries related to each aspect rating. These queries were based on real queries made by users in popular search engines, so they reflect a natural user query. For example, the query “great location” is related to the aspect rating *location*. Given the query, the dataset lists the aspect rating value of each hotel as described in Table 1. The rating values are given by users from TripAdvisor when evaluating the hotels they have visited. In essence, the OpinRank dataset contains five hotels aspects, each aspect is related to five user queries and one aspect rating value for each hotel as described in Table 1.

4.2 Methodology

The DBpedia and LinkedGeoData were accessed through the local repository, or by the Snorql endpoint, as explained in Subsection 4.1. All experiments were executed in the same computer with an Intel Processor of 1.8 GHz (model i3-3217U) and 8 GB of RAM memory. For processing the SKPQ we made use of OpenStreetMap dataset, while for SKPQ-LD we used DBpedia dataset merged with OpenStreetMap dataset using SPARQL queries.

The experiments were employed with two methodologies to evaluate the SKPQ-LD: using ratings obtained from Google Places API, and relevance judgments obtained from TripAdvisor. In Experiment 1, we apply the first methodology, where SKPQ and SKPQ-LD were executed twenty times using one unique query keyword each time. Half of the keywords are the most frequent terms in

the dataset, the other half were randomly obtained. The query results were evaluated using NDCG. The list of frequent terms was obtained from S2I⁴ and random queries keywords were obtained without repetition from a set of 1906 terms extracted from the OpenStreetMap dataset. “chili” and “sunset” are examples of random keywords used in this work. We used the object rate obtained from Google Places API to determine the ideal ranking.

In Experiment 2, we apply the second methodology, where SKPQ and SKPQ-LD were executed using query keywords described in the OpinRank dataset. This dataset contains full reviews of hotels collected from Tripadvisor and their corresponding aspect ratings as described in Subsection 4.1. We use the queries related to each aspect as query keywords and evaluate the query result obtained by SKPQ and SKPQ-LD. We ordered the query result by the aspect rating value of each hotel to determine the ideal ranking.

The metrics employed in all experiments were Discount Cumulative Gain (DCG), Normalized Discount Cumulative Gain (NDCG) and Mean Average Precision (MAP). The NDCG is widely used in Information Retrieval, measuring the quality of the ranking produced by a system [10]. It is particularly suitable for search applications since it accounts for multilevel relevance. The NDCG corresponds to the value of DCG divided by IDCG, defined in Equation 3. Since the top-k items are presented in a rank, then the Discounted Cumulative Gain at position k (DCG@k) and ideal DCG (IDCG) are calculated based on Equation 1 and 2, respectively. We denote top-k items by $P_k = \{p_1, p_2, \dots, p_k\}$, where the items are ranked by the SKPQ and SKPQ-LD; and we denote rel_i as the relevance value of the item at position i .

$$DCG@k = \sum_{i=1}^{|P_k|} \frac{rel_i}{\log_2(i+1)} \quad (1)$$

The IDCG is the maximum value of DCG. It is calculated as

$$IDCG = \max(DCG@k) \quad (2)$$

So, NDCG@k is calculated as

$$NDCG@k = \frac{DCG@k}{IDCG} \quad (3)$$

4.3 Experiment 1: Evaluating Query Results

To understand the ranking quality of both SKPQ and SKPQ-LD, we compared the NDCG values obtained when using random keywords and frequent keywords. Figure 3 reports the arithmetic mean of NDCG@k (k=5, 10, 15, 20) that are generated by the queries with different keywords. The arithmetic mean values are reported on the vertical axis. Figures 2(a) and 3(a) illustrate that SKPQ-LD improves the ranking quality when using random keywords, otherwise the quality is roughly the same.

It is noticeable that we obtain better results with SKPQ using frequent keywords. Since the keyword is present in many objects, there is no problem to SKPQ identify the object that has textual relevance to the query keyword. In this scenario, the objects in

⁴Implementation available at XXL Library

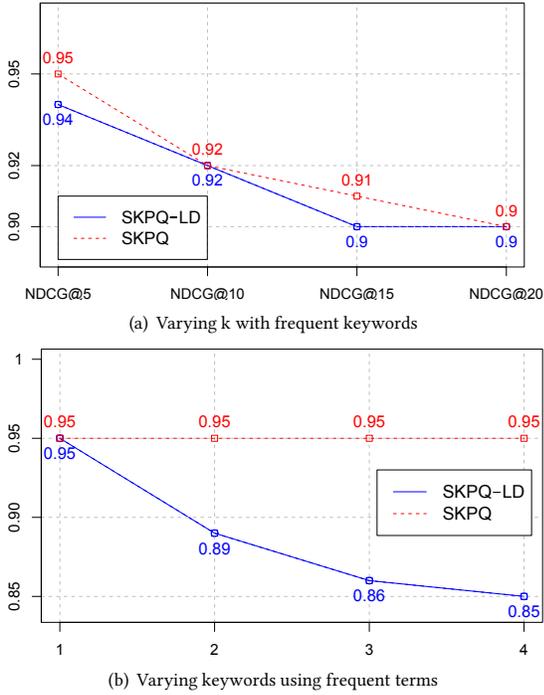


Figure 2: Results obtained by SKPQ and SKPQ-LD varying the keywords and the query result size (k)

SKPQ have a small textual description, but they have a high probability to match with the query keyword. In addition, the SKPQ access more objects because OpenStreetMap offers a larger dataset. Therefore, SKPQ counts on a good enough textual description, and a larger amount of objects, factors that lead to a better evaluation result. Nevertheless, the SKPQ-LD obtained results nearly as good as SKPQ, with a difference of only 0.1 between the NDCG values.

Figures 2(b) and 3(b) illustrate the NDCG values obtained when varying the number of query keywords. The results depicted in this Figure use a fixed k value of 5. The experiment illustrated in Figure 2(b) used the 10 most frequent terms in the dataset as query keywords. To build query keywords with 2 terms or more, we combined these terms with each other without repetition.

As it can be seen in Figure 2(b), even after adding three more keywords, the results obtained in SKPQ does not change. On the other hand, SKPQ-LD is more influenced by the increase in the number of query keywords. As observed in Figure 3, the SKPQ presents better outcomes with frequent keywords while SKPQ-LD is better with random keywords. However, the distance between NDCG values obtained by SKPQ-LD in Figure 2(b) slowly decreases as the number of keywords grows. In addition, we noticed that the SKPQ results had few, or none, changes when the number of keywords was increased. For example, the query result for the keywords “parking cafe” was equal to the query results obtained with “bank parking cafe” and “parking supermarket cafe bank”. The textual score of each object presented had changed, but there was no difference on the rank order, resulting in similar

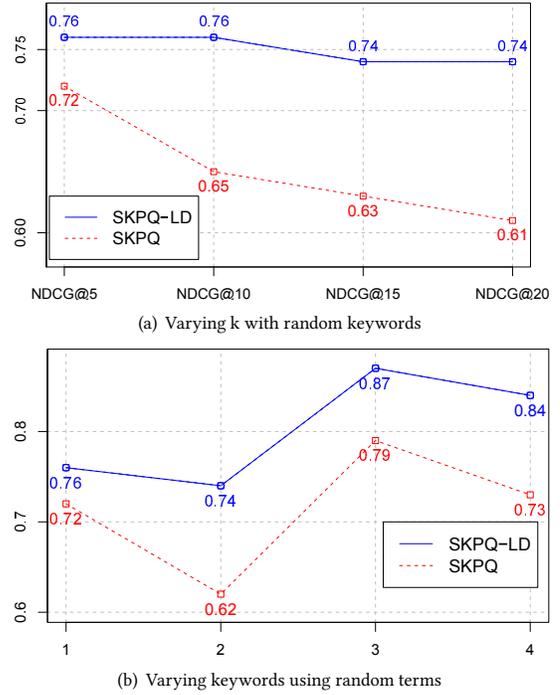


Figure 3: Results obtained by SKPQ and SKPQ-LD varying the keywords and the query result size (k)

NDCG values. The SKPQ lacks a result variability because of the poor textual description of its objects. SKPQ-LD obtained lower NDCG values but did present different results to each query keyword.

As a baseline, the SKPQ query results are compared against the top- k Range Query (RQ) [3] results. We employ our approach to enrich the textual description of objects accessed by RQ and evaluate the results obtained. Given a spatial area and the query keyword, the RQ returns k objects in the given area that are textual relevant to the query keyword. All RQ used the same query keywords as SKPQ and a random query location in Dubai. The radius of 200 m from the selected query location defines the spatial neighborhood.

It can be seen in Figure 4 that our approach improved RQ result set when using frequent keywords instead of random keywords. The RQ looks for all k objects in a small spatial area (radius = 200 m) while SKPQ looks for objects in the neighborhood of many objects of interest. Each object neighborhood has the same size of all the spatial area visited by RQ (200 m). This contrast results in a more challenging effort to build a quality rank for the given area because there are fewer objects to verify. This can be verified observing the much lower NDCG values obtained with RQ. While SKPQ obtained 0.61 in its worst case, RQ obtained 0.41 as its best case. The amount of objects to verify is the main reason for the lower NDCGs values depicted in Figure 4 than the ones in Figure 3.

Figure 5 illustrates the relative NDCG improvement (as described in [13]) of the proposed approach e_{pro} over respective baseline model e_{other} , further measured as

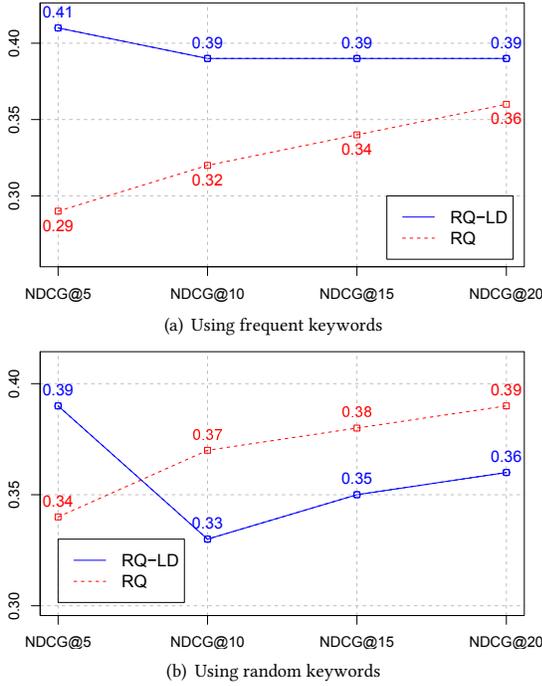


Figure 4: Results obtained with RQ and RQ-LD

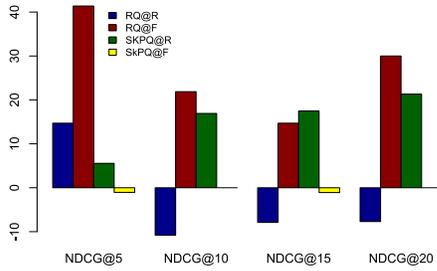


Figure 5: Relative NDCG improvements

$$(e_{pro} - e_{other}) / e_{other} \times 100 \quad (4)$$

Figure 5 reports the relative NDCG improvement values on the vertical axis. The proposed approach demonstrated different degrees of improvement in different scenarios. It improved SKPQ relative NDCG in 20% when using random keywords (SKPQ@R - NDCG@20) and 40% when RQ used frequent keywords (RQ@F - NDCG@5).

Using the users' ratings obtained from Google Maps, we evaluate if our approach improves the query result. Using random keywords, the hotels presented as query results on SKPQ-LD are more popular among the users than the ones presented by the SKPQ. Using frequent keywords, the query result quality on SKPQ-LD is very similar to the one obtained by the SKPQ. Therefore, our approach does not impose a high penalty over the quality of the query result.

4.4 Experiment 2: Evaluating feature selection

In Experiment 2, we used the queries in OpinRank to evaluate the feature selection in SKPQ and SKPQ-LD. Since the OpinRank dataset contains only hotel reviews, we restrict our feature dataset to hotels. All hotels used in this experiment are located in Dubai.

Given the query keywords, the SKPQ returns a list of objects of interest whose are near to features and are relevant to the given query. We desire that SKPQ returns objects whose features have a high aspect rating value. This way, the SKPQ would be selecting good features according to users of TripAdvisor. If there is no relevant feature near an object of interest, the SKPQ query result is empty.

The OpinRank dataset offers 5 textual queries for each aspect rating (total of 25 queries). These textual queries were used as query keywords in SKPQ. However, SKPQ did not find any feature whose textual description was relevant to the query keywords. The description used in SKPQ was too short and could not describe the feature as needed. Notwithstanding, the SKPQ-LD was able to find textual relevant features. From 25 queries, SKPQ-LD was able to find relevant features in 15 (equals to 60% of all executed queries). The features were retrieved with different degrees of textual relevance. Considering $k = 5$ and 25 as the number of executed queries, the MAP score obtained was 0.46.

Between the 15 relevant query results obtained by SKPQ-LD, we could extract the aspect rating value of few features. Many times, the hotel name in OpinRank dataset was not found in DBPedia or OpenStreetMap. Hence, when SKPQ or SKPQ-LD returns a hotel name that does not appear in the OpinRank dataset we can not retrieve its aspect rating value.

We show examples of textual queries that we could extract rating values, and those we could not, to illustrate this scenario. The queries "nice staff" and "good value" are examples of queries that did not return any relevant objects to the user. The objects textual description in SKPQ and SKPQ-LD was not able to describe these aspects of the hotels. However, the queries "great location", "clean place" and "cozy rooms" returned objects when using SKPQ-LD. Figure 6 reports the NDCG values of the query results obtained with these query keywords.

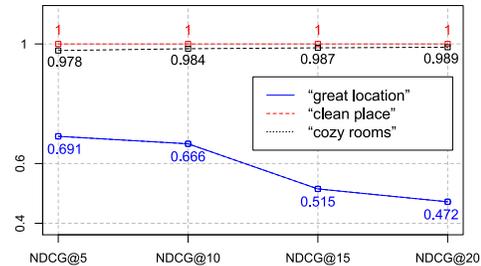


Figure 6: SKPQ-LD evaluation using OpinRank

With the enhancing of objects' textual description, SKPQ-LD was able to select more objects that satisfy the user need than SKPQ. Accordingly to the obtained NDCG values in Figure 6, SKPQ-LD selected features of good quality. Since the query results have high aspect rating values, we can assume that SKPQ-LD was

able to find good objects to the user. For the query “clean place” for example, SKPQ-LD was able to find features that are evaluated by real users as a clean hotel.

The OpinRank dataset contains other queries created by the combination of the queries illustrated in Figure 6 plus the queries “nice staff” and “good value”. Nevertheless, the combination of these queries lead to results very similar to the ones at Figure 6. In this experiment, the SKPQ-LD demonstrated that the textual description improvement enhances the query capabilities, enabling it to find more objects. Without the textual description improvement, the SKPQ was unable to find any relevant objects to the presented queries.

5 LIMITATIONS AND POINTS OF IMPROVEMENTS

Despite the obtained results look promising, our approach has some limitations. First, although the LOD cloud increases every day, textual descriptions may not always be available with expected quality. This may eventually penalize the query results when using LOD.

Zarrinkalam and Kahani [14] describe an enrichment approach using LOD to improve the textual description of articles citations. Accordingly to him, “the Linked Data driven enrichment process has improved the quality of recommendations but it isn’t as much as expected” because of “data sources that publish bibliographic information on the LOD cloud, do not yet provide adequately rich and high-quality data, compared to what these data sources provide on the web of documents”.

We face the same problem with spatial information on LOD objects. LinkedGeoData has a higher amount of objects registered than DBpedia. But the textual description of objects in LinkedGeoData is poor as the ones in OpenStreetMap. In addition, a lot of less popular objects are not registered on DBpedia yet or are not well documented. Many objects do not have the *geo:geometry* property too. As a consequence, the textual description of some objects can not be enriched. For this reason, the results obtained by our approach is lower than the ones obtained by the traditional SKPQ when using frequent keywords in Experiment 1. Since the term used as the keyword is frequent in the OpenStreetMap dataset, there is no need for textual description enrichment. If we are looking for objects described as “restaurant” and all restaurants are described in the dataset, there is no need for a more detailed description. The SKPQ performs better in this context because its objects have the description needed and it has access to more objects, so it can search for more restaurants that satisfy the user need.

The world of Linked Data poses many challenges, as described in [8] and [2]. One meaningful challenge is the data integration in the complex and schema-less Semantic Web. However, with the fast growth of the LOD cloud, the semantic annotation becomes more popular and the datasets will provide more quality data. The proposed approach will be even more effective when more high quality data becomes more present in the web of data.

6 CONCLUSION

In this paper, we proposed an enhancement to Top-k Spatial Keyword Preference Query. This enhancement uses LOD to improve

the textual description of features. Results from our experiments show that a richer textual description can contribute to enhancing the SKPQ query result.

A larger textual description was employed to present results for the user in situations where traditional SKPQ could not. In the first experiment conducted, evaluating the query results with Google Maps dataset, we observed that our method can perform 20% better than the traditional approach. This takes place because all objects had a wider textual description when processing the query using our approach. Also, in Experiment 2 we observed that using real queries obtained from OpinRank dataset, the SKPQ was unable to find features without using our approach. In addition to finding these features, we observed that the selected features have good quality according to TripAdvisor users.

In future works, we aim at creating an evaluation model using expertise judgments. This will give a more precise evaluation about the SKPQ-LD. We also plan to extend the algorithm, enabling richer textual descriptions. Moreover, we also have the intention to evaluate the response time and I/O of the SKPQ-LD. These measures will be useful to analyze the impact of LOD on query processing performance.

Acknowledgments. We thank FAPESB for the financial support of this research.

REFERENCES

- [1] Christian Becker and Christian Bizer. 2009. Exploring the geospatial semantic web with dbpedia mobile. *Web Semantics: Science, Services and Agents on the World Wide Web* (2009), 278–286.
- [2] Christian Bizer, Peter Boncz, Michael L Brodie, and Orri Erling. 2012. The meaningful use of big data: four perspectives—four challenges. *ACM SIGMOD Record* 40, 4 (2012), 56–60.
- [3] Xin Cao, Lisi Chen, Gao Cong, Christian S. Jensen, Qiang Qu, Anders Skovsgaard, Dingming Wu, and Man Lung Yiu. 2012. Spatial keyword querying. In *ER*. Springer, 16–29.
- [4] Gao Cong, Christian S. Jensen, and Dingming Wu. 2009. Efficient retrieval of the top-k most relevant spatial web objects. *PVLDB* 2, 1, 337–348.
- [5] João Paulo Dias de Almeida and João B Rocha-Junior. 2016. Top-k Spatial Keyword Preference Query. *Journal of Information and Data Management* (2016), 162.
- [6] Ignacio Fernández-Tobías, Iván Cantador, Marius Kaminskas, and Francesco Ricci. 2011. A generic semantic-based framework for cross-domain recommendation. In *Proceedings of the 2nd International Workshop on Information Heterogeneity and Fusion in Recommender Systems*. ACM, 25–32.
- [7] Kavita Ganesan and ChengXiang Zhai. 2011. Opinion-Based Entity Ranking. *Information Retrieval* (2011).
- [8] Jorge Gracia, Elena Montiel-Ponsoda, Philipp Cimiano, Asunción Gómez-Pérez, Paul Buitelaar, and John McCrae. 2012. Challenges for the multilingual web of data. *Web Semantics: Science, Services and Agents on the World Wide Web* 11 (2012), 63–71.
- [9] Vinod Hegde, Vinny Reynolds, Josiane Xavier Parreira, and Manfred Hauswirth. 2011. Utilising Linked Data for Personalized Recommendation of POI’s. In *International AR Standards Meeting, Barcelona, Spain*.
- [10] Kalervo Järvelin and Jaana Kekäläinen. 2002. Cumulated gain-based evaluation of IR techniques. *ACM Transactions on Information Systems (TOIS)* 20, 4 (2002), 422–446.
- [11] Roula Karam and Michele Melchiori. 2013. Improving geo-spatial linked data with the wisdom of the crowds. In *Proceedings of the joint EDBT/ICDT 2013 workshops*. ACM, 68–74.
- [12] Matthew Perry and John Herring. 2012. OGC GeoSPARQL-A geographic query language for RDF data. *OGC Implementation Standard*. Sept (2012).
- [13] Hengjie Song, Yonghui Xu, Huaqing Min, Qingyao Wu, Wei Wei, Jianshu Weng, Xiaogang Han, Qiang Yang, Jialiang Shi, Jiaqian Gu, et al. 2016. Individual Judgments Versus Consensus: Estimating Query-URL Relevance. *ACM Transactions on the Web (TWEB)* 10, 1 (2016), 3.
- [14] Fattane Zarrinkalam and Mohsen Kahani. 2012. A multi-criteria hybrid citation recommendation system based on linked data. In *Computer and Knowledge Engineering (ICCKE), 2012 2nd International eConference on*. IEEE, 283–288.

Kumo: um serviço para portabilidade em multi-nuvens heterogêneas

Marcus Rafael Xavier
Universidade Federal de Pernambuco
Jornalista Anibal Fernandes, s/n,
Cidade Universitária
Recife, Pernambuco 50740-560
mrxl@cin.ufpe.br

Carlos Andre G. Ferraz
Universidade Federal de Pernambuco
Jornalista Anibal Fernandes, s/n,
Cidade Universitária
Recife, Pernambuco 50740-560
cagf@cin.ufpe.br

Ioram Schechtman Sette
Centro de Estudos e Sistemas
Avançados do Recife (CESAR)
Rua Bione, 220, Bairro do Recife
Recife, Pernambuco 50030-390
iss@cesar.org.br

ABSTRACT

Cloud computing is maturing and becoming ubiquitous in people's daily lives. As a result, cloud environments are providing more and more services with better quality of service. Cloud customers, however, have suffered from the vendor lock-in problem, in such a way that those who wish to migrate to another cloud provider require partial or total reimplementation of applications and virtual infrastructure. Moreover, the problem of heterogeneity among distinct cloud environments makes it difficult for the portability of resources between them. Therefore, this work focuses on the development of an ontology to handle multi-cloud heterogeneity, and thus, bring interoperability in the form of a service to perform the portability of virtual machines between different providers.

KEYWORDS

Multi-nuvem, Aprisionamento, Heterogeneidade, Portabilidade.

1 INTRODUÇÃO

Com o passar dos anos, várias tecnologias como virtualização, computação em grade, e arquiteturas orientadas a serviço (SOA) têm amadurecido e contribuído significativamente para tornar a computação na nuvem viável [9].

A computação na nuvem permitiu que os serviços web (do inglês, *web services*) fossem ainda mais disseminados, possibilitando a hospedagem de aplicações web em larga escala, e fazendo com que empresas dependam cada vez mais de recursos em nuvem para implantação de seus sistemas.

Juntamente com o atual amadurecimento e migração de infraestruturas para as nuvens, novos problemas surgem. No cenário com uma única nuvem destaca-se o aprisionamento tecnológico ou *vendor lock-in*, enquanto que no cenário multi-nuvem, cenário cujo o aprisionamento tecnológico não é presente, destaca-se o problema da heterogeneidade.

A solução pra esses e outros problemas da computação na nuvem deve-se começar por estratégias que tragam interoperabilidade para esses ambientes heterogêneos, permitindo assim que usuários se libertem do aprisionamento causado pelas implementações proprietárias e não padronizadas promovidas pelos provedores de serviços de nuvem.

Inicialmente, é necessário o uso de alguma estratégia para que se possa conhecer as ações comuns a todas as nuvens, as comuns apenas a algumas nuvens, e ações que são específicas de cada nuvem (ver Figura 1). Logo, por ser indicada para representação formal de conhecimento, através de classes (*server*), relacionamento (*tem um ou vários*), e atributos (*volume*), optou-se pelo uso de uma ontologia.

Neste trabalho é realizado o desenvolvimento de uma ontologia com foco em portabilidade de máquinas virtuais. Com isso, busca-se conhecer as operações sobre essas VMs de modo que se possa a partir dessa ontologia, desenvolver um serviço que possa realizar o gerenciamento de recursos, como máquinas virtuais, na estratégia de multi-nuvem heterogêneas.

Entretanto, a principal motivação é a especificação de um sistema chamado Kumo, que foca em realizar migração de máquinas virtuais entre multi-nuvens heterogêneas. As migrações de máquinas virtuais são atualmente um problema em multi-nuvem em ambientes homogêneos e heterogêneos, e esse é o desafio que esse trabalho trata.

2 CONTEXTUALIZAÇÃO

Essa seção é dedicada a contextualização dos problemas principais em ambientes de múltiplas nuvens, o de aprisionamento tecnológico, também conhecido na literatura como *vendor lock-in*, e a heterogeneidade das tecnologias e interfaces implementadas pelos provedores de serviços de nuvem (CSPs).

A computação em nuvem tornou-se uma tendência para empresas construírem suas infraestruturas virtuais, em vez de usar recursos locais (do inglês, *on-premises*). Com isso, essas empresas que utilizavam recursos locais iniciaram a migração para a nuvem, e aquelas novas empresas criadas a partir de então, já utilizam a nuvem como a plataforma para receber toda sua infraestrutura.

Com isso, um problema relevante na computação na nuvem surge. Uma vez que uma empresa estabelece sua infraestrutura em um provedor de nuvem, será difícil ou até mesmo impossível a migração para outro provedor. Sendo necessária a reimplementação parcial ou total da infraestrutura, bem como de aplicações desenvolvidas utilizando bibliotecas fornecidas pelo provedor, que são compatíveis apenas com seus serviços.

Quando essa situação acontece, é dito que uma empresa está aprisionada tecnologicamente ao seu provedor de nuvem, esse que não fornece opções na forma de ferramentas para facilitarem a migração para um outro provedor [3]. Assim, uma possibilidade não otimizada é a de baixar os dados e configurações feitas, para que possam ser importados e replicados em outro provedor. Porém

mesmo nesse caso mais manual de migração, um segundo problema pode acontecer, esse problema é a heterogeneidade.

No contexto da portabilidade de recursos entre nuvens, heterogeneidade é compreendida como a situação em que um usuário que é detentor de um recurso em um dado formato de arquivo proprietário, e que esse formato só pode ser compreendido pelo provedor que o usuário ou empresa está aprisionado. Com isso, mesmo de posse do arquivo que representa o recurso que é necessário migrar, não há garantia que o provedor de destino possa entender e executar esse recurso.

3 INTEROPERABILIDADE

Para propor qualquer solução para o problema de aprisionamento tecnológico, um bom ponto de partida é entender o que significa a ideia de trazer interoperabilidade para as nuvens. De acordo com o documento da organização Object Management Group [4], interoperabilidade *"deve ser vista como a capacidade de serviços de nuvem pública, serviços de nuvem privada e outros sistemas diversos dentro da empresa de entender as interfaces de aplicativos e serviços de cada um, a configuração, as formas de autenticação e autorização, formatos de dados, etc. a fim de trabalhar uns com os outros."*

Na computação na nuvem, interoperabilidade está diretamente relacionada a compatibilidade entre os tipos e formatos de dados usados para comunicação entre interfaces de programação (APIs) como o JSON¹, entre linguagens de programação e aplicações que possam ser executadas por vários provedores. Também se pode considerar interoperabilidade entre componentes de infraestrutura, como arquivos referentes a máquinas virtuais (arquivos de *appliances*) executáveis por vários hipervisores, ou imagens de contêineres.

Sendo a interoperabilidade uma característica fundamental não apenas em ambientes de nuvem, mas também nos ambientes multi-nuvem, sejam eles homogêneos ou heterogêneos. Ainda existem muitos desafios para instalação, manutenção e operação desses ambientes pela falta dessa interoperabilidade. Em [9], oito desses desafios estão listados e aqui descritos quanto a sua motivação para ambientes de multi-nuvem.

1. **Rede** - Conectividade (gerência da infraestrutura física ou virtual de rede); Endereçamento (mobilidade do endereço IP); *Naming* (nomes para identificar recursos nas várias nuvens); *Multicasting* (envio seletivo de pacotes para dispositivos em nuvens diferentes).
2. **Economia** - Mercado (estratégias de negociação referente ao uso recursos de outras nuvens); Preço (definições de como comprar/vender/alugar/emprestar recursos com outras nuvens); Contabilidade e Faturamento (quantificação de uso de recursos de/por terceiros e estratégias de pagamento, pré e pós-pago).
3. **Provisionamento** - Descoberta (descoberta de recursos em outras nuvens); Seleção (escolha por provedores baseado nas garantias e qualidade de serviço); Alocação (estratégias de reservar recursos previamente ou sob demanda em outras nuvens).
4. **Acordo de Nível de Serviço (SLA)** - Gerência (gerenciamento de contratos SLA pelas nuvens); Acordo de Nível de Serviço da Federação (contrato que determina quais

condições um nuvem precisa para interoperar com outras nuvens); Monitoramento e Dependência (verificação das nuvens quanto o cumprimento dos contratos de SLA).

5. **Segurança** - Confiança (garantia de confiança baseado em ranking de provedores mais confiáveis); Autorização e Autenticação (acesso e ações que um nuvem poderá usar de outra quando previamente autenticada); Políticas e Interoperabilidade Semântica (garantia de equivalência semântica quanto às permissões das nuvens).
6. **Autonomia** - gerência autônoma dos recursos nas múltiplas nuvens.
7. **Monitoramento** - garantia de coleta de métricas de utilização em múltiplas nuvens.
8. **Portabilidade** - Portabilidade de Dados (padronização de dados e de seus formatos); Portabilidade de VMs (migração de VMs entre nuvens homogêneas e heterogêneas).

3.1 Estratégias

Em [9] são mencionadas quatro alternativas para trazer interoperabilidade para nuvens, divididas em duas categorias. Primeiramente são apresentadas duas categorias onde a mudança necessária para alcançar esse requisito é implementada e centrada nos provedores de nuvem (*provider-centric*), e em seguida são apresentadas outras duas abordagens, centradas e implementadas nos clientes de nuvem (*client-centric*).

3.1.1 Nuvem híbrida. A interoperabilidade de nuvem híbrida é uma abordagem *provider-centric* que permite que um provedor de serviço tenha uma parceria com uma nuvem pública. Assim que a parceria for formada, se o uso de recursos da nuvem privada aumentar e não houverem mais recursos disponíveis, a nuvem privada utiliza recursos do provedor de serviços de nuvem pública.

3.1.2 Nuvem federada. A federação de nuvem é outra abordagem *provider-centric* para fornecer interoperabilidade na computação em nuvem. No cenário de federação, os provedores de nuvem compartilham seus recursos ociosos dos *data centers*, com os regulamentos e o acordo de nível de serviço (SLA) definido pelos membros da federação.

3.1.3 Aplicações multi-nuvem. Aplicativos com várias nuvens são uma abordagem de interoperabilidade *client-centric* que permite aos desenvolvedores escreverem suas aplicações independente da tecnologia de nuvem. Se necessário migrar para outro provedor, o aplicativo deve estar pronto para ser reimplantado sem muito esforço no novo ambiente. Os *frameworks* e bibliotecas como o projeto Apache Libcloud² podem ser usados para auxiliar esse desenvolvimento.

3.1.4 Broker multi-nuvem. O broker multi-nuvem é outra abordagem de interoperabilidade *client-centric*. Nela, as requisições feitas pelos usuários são recebidas pelo broker, que cria uma nova requisição com a mesma semântica para a nuvem do provedor de serviço escolhida pelo usuário.

¹www.json.org

²libcloud.apache.org

4 TRABALHOS RELACIONADOS

Na presente literatura são documentados trabalhos relacionados a identificação de formas para conquistar a interoperabilidade em ambientes multi-nuvs. No sentido de migração de aplicações entre provedores, uma proposta encontrada em [7] foca na combinação da arquitetura de micro-serviços para a separação de responsabilidades, e na utilização de contêineres para empacotamento das aplicações e de suas dependências.

Na estratégia de utilização de contêineres, é possível portar as aplicações devido vários provedores utilizarem as mesmas tecnologias para a execução e orquestração desses contêineres. O Docker³ é atualmente a plataforma para contêineres mais difundida entre os provedores, enquanto que a ferramenta de orquestração Kubernetes⁴ é a mais adotada para realizar interoperabilidade de contêineres Docker, permitindo migração dos contêineres entre infraestruturas homogêneas (que executem Docker).

Diferentemente, em [5] é descrita uma outra alternativa para interoperabilidade multi-nuvem. Nessa abordagem, usuários administradores da plataforma *LambdaLink* podem cadastrar imagens de discos virtuais que podem ser utilizadas para o lançamento de instâncias em múltiplas nuvs heterogêneas. Entretanto, o problema da portabilidade de máquinas virtuais entre as nuvs não é tratado pela plataforma. Em resumo, tal plataforma permite apenas o compartilhamento de imagens de disco.

Diferentemente de [7] e [5], este trabalho trata o problema de máquinas virtuais criadas em ambientes de nuvem, possibilitando a portabilidade das VMs entre multi-nuvs de tecnologias heterogêneas. Mesmo com propósito diferente, a proposta [7] se mostra promissora para a implantação do serviço Kumo, que pode ser implementado utilizando micro-serviços (permite a evolução do Kumo para agregar novas funcionalidades, como o gerenciamento multi-nuvem) quanto o uso contêineres para implantação do Kumo, trazendo a possibilidade de migrar o serviço e ter múltiplas instâncias do mesmo. Já como no trabalho [5], Kumo poderia ter um repositório de imagens de disco para facilitar além da migração, a criação de novas máquinas virtuais.

5 O PROBLEMA DA PORTABILIDADE

Como mencionado, a portabilidade é um desafio na computação em nuvem. O domínio de alguns provedores de nuvem em relação a outros aumenta o aprisionamento de clientes a uma única plataforma de nuvem, impedindo a portabilidade de dados ou software criado por eles [1]. Nesta seção, algumas motivações e desafios relevantes em relação à portabilidade são discutidos.

Em [1], alguns pontos interessantes são discutidos sobre a comunidade de computação em nuvem. Em primeiro lugar, a comunidade deve promover padrões comuns e interoperabilidade entre os serviços de nuvem pública, o que maximizaria os ganhos econômicos para os provedores e também seria um ponto de partida para transformar a portabilidade multi-nuvem em realidade. Em seguida, a comunidade precisaria desenvolver a mentalidade para divulgar amplamente o princípio de facilitação do acesso a dados para conquistar a portabilidade, evitando assim o aprisionamento.

³www.docker.com

⁴kubernetes.io

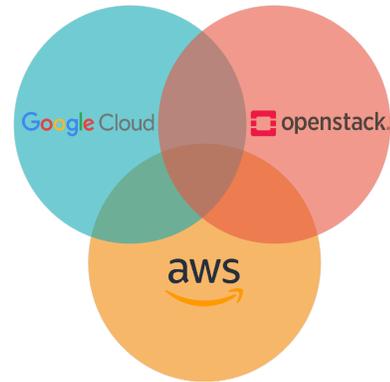


Figura 1: Nuvs selecionadas para a ontologia

Um exemplo disso é a comunidade OpenStack⁵, que é um projeto de código aberto e gratuito para lançar nuvs privadas e públicas. Embora o OpenStack seja desenvolvido por qualquer pessoa que queira contribuir e que muitos colaboradores do projeto sejam funcionários dos provedores de nuvem, a comunidade garante que projetos de código aberto e até mesmo fechados criados por fornecedores, interoperem com o OpenStack.

As motivações para a portabilidade são, além da eliminação do aprisionamento, a possibilidade de importar e exportar dados (por exemplo, motivado por uma outra nuvem que tenha uma estratégia de recuperação de desastres melhor); SLA aprimorado (por exemplo, um recurso pode ser executado melhor no ambiente de nuvem de outro provedor com um hipervisor de melhor desempenho para execução de uma máquina virtual) e também a opção de escolha baseada no preço (outra nuvem pode ter um data center em uma região com eletricidade mais barata e assim permitir menor preço [3]).

Tão importante quanto migrar os dados de uma nuvem para outra, é migrar componentes de infraestrutura para outra nuvem, por exemplo, para satisfazer aos requisitos tecnológicos de uma aplicação [1]. A migração de uma máquina virtual de uma nuvem heterogênea envolve muitas decisões, por exemplo, decidir se a migração será executada de forma *live* (neste caso, o usuário continua a usar a máquina virtual original até que uma cópia dela seja lançada na nuvem de destino); se o hipervisor da máquina virtual na nuvem de destino pode executá-la; decidir como configurar a rede; o que fazer se a máquina virtual tiver um volume anexado; e assim por diante.

Nesse sentido, considerando uma migração feita de forma manual por um operador de nuvem, para cada passo necessário durante a migração, esse operador deverá conhecer todos os comandos necessários para a execução da migração, isso para cada uma das tecnologias das nuvs heterogêneas, iniciando em como autenticar em todas as nuvs envolvidas na operação. Para lidar com tais preocupações, a literatura mostra que o conceito de ontologia pode ser usado para mapear recursos às ações possíveis ao recurso "máquina virtual". Na seção 6 essa abordagem é discutida e desenvolvida.

6 ONTOLOGIA

Conforme mencionado em [1], no contexto da computação em nuvem, uma ontologia pode ser usada para obter interoperabilidade entre diferentes provedores de nuvem e seus serviços.

O conhecido trabalho "*O que são ontologias, e por que precisamos delas?*" [2], define o que é uma ontologia:

Ontologia é um vocabulário de representação, geralmente especializado em algum domínio ou assunto. Mais precisamente, não é o vocabulário como tal que se qualifica como uma ontologia, mas os conceitos que os termos do vocabulário pretendem captar. Assim, traduzir os termos de uma ontologia de uma língua para outra, por exemplo, do inglês para o francês, não altera conceitualmente a ontologia.

Devido a heterogeneidade das nuvens, se houver necessidade de portabilidade multi-nuvem, uma ontologia pode ser um recurso para entender o que é comum a todas as nuvens, o que é comum em algumas nuvens, e o que é específico de cada nuvem. Com isso, é possível conhecer onde a heterogeneidade nas nuvens acontece e também definir um "*vocabulário especializado*" comum que, parafraseando [2], permite traduzir os termos da ontologia de uma nuvem para outra, por exemplo, da Amazon Web Services para OpenStack, sem conceitualmente alterar a ontologia.

Na criação de uma ontologia para interoperabilidade, primeiramente é fundamental definir quais componentes dentre os vários disponíveis na computação em nuvem, são relevantes para a portabilidade das máquinas virtuais. Usando a experiência da utilização da nuvem, entende-se que a máquina virtual é um recurso gerenciado pelo serviço de computação. No entanto, uma observação que vem à mente é que a máquina virtual tem relação não só com o serviço de computação (*computing*), mas está relacionada ao serviço de rede (*networking*), imagem de disco (*image*), e discos virtuais (*volume*).

Para a ontologia desenvolvida neste trabalho, foi utilizada uma estratégia *bottom-up* onde as nuvens foram analisadas a partir de suas definições de políticas de autorização, para que em seguida fosse possível encontrar aqueles recursos que são comuns as políticas de cada uma das nuvens, e assim mapear as ações comuns a cada um desses recursos. Esse trabalho foca exclusivamente em recursos e em ações comuns às três nuvens, que é representado pela interseção entre os conjuntos na Figura 1.

Dessa forma, a primeira decisão tomada para iniciar o desenvolvimento da ontologia, foi estabelecer que o recurso foco da ontologia seria a máquina virtual (*server*, na ontologia), não apenas considerando como parte do serviço de computação, mas como parte de todos os serviços que são relacionados a ela, como rede, imagem de disco e volume. Por exemplo, do componente de armazenamento em bloco foi considerado o recurso volume, isso porque a máquina virtual pode ter um volume associado. Outro exemplo ocorre no serviço de rede, onde foi considerada a interface de rede, pois este recurso associado a máquina virtual para trazer comunicação a ela via rede. Em resumo, não foram adicionados à investigação recursos de serviços que não têm relação nenhuma com o recurso principal da ontologia, a máquina virtual.

⁵www.openstack.org/community

Dando continuidade ao desenvolvimento da primeira versão da ontologia, foram escolhidas as nuvens que seriam usadas nesse trabalho, que foram:

Amazon Web Services (AWS) é uma plataforma de nuvem pública que oferece computação, armazenamento de banco de dados, entrega de conteúdo e etc. Atualmente, é o serviço de nuvem pública mais usado [8].

Google Cloud Platform (GCP) é a opção de serviço de nuvem pública da Google. Ela fornece menos serviços do que a AWS, mas por outro lado, tem sido a primeira opção dos usuários de nuvem quando pensam em migrar de seu provedor atual para um novo provedor [8].

OpenStack (OS) é um software de código aberto para implantação de nuvens públicas e privadas, capaz de gerenciar recursos de computação, armazenamento e rede [6].

Após a escolha pelas três nuvens, a etapa seguinte foi definir quais informações seriam usadas para extrair os recursos e suas ações. Para isso, chegou-se as políticas de autorização do Amazon Web Services ⁶ (AWS), do Google Cloud Platform ⁷ (GCP) e do OpenStack ⁸ (OS). A escolha por essas políticas veio por serem recursos estruturados e que contemplam todas as possibilidades em ações que um usuário, com o papel de administrador pode realizar em cada uma dessas nuvens.

Para encontrar as políticas na AWS foi necessário fazer login com uma conta, navegar para o Identity and Access Management (IAM) e acessar as políticas para o serviço de computação EC2, para o papel denominado *AmazonEC2FullAccess*. As políticas do GCP, análogas à AWS, foram encontradas no seu serviço de IAM, onde as políticas de autorização para o papel *Compute Admin* puderam ser encontradas. No OpenStack, foi necessário acessar a página do GitHub⁹ para todos os serviços necessários: o Nova (computação), o Cinder (volume), o Neutron (rede) e o Glance (imagem).

Após a criação de uma tabela para receber todas as políticas de autorização para as três nuvens, as documentações dessas nuvens também foram analisadas, uma a uma. O objetivo dessa etapa foi descobrir possíveis ações que tivessem sintaxe diferentes, mas que semanticamente representassem a mesma ação. Conceitualmente, o resultado do trabalho até este ponto é representado pela interseção entre os três conjuntos na Figura 1.

Em seguida foi criado um mapeamento entre os recursos e suas ações para as três nuvens. Um fragmento desse mapeamento pode ser visto na Tabela 1 que representa o recurso *Image* (imagem) nas três nuvens. Assim, é possível ver como foram mapeadas cada uma das ações de *Image*, de maneira que todas as ações da ontologia (última coluna da Tabela 1) sejam nomes em alto nível e representem semanticamente o que cada uma das ações significa.

O resultado conquistado na ontologia pode ser conferido na Figura 2, que representa a ontologia na linguagem *Ontology Web Language* (OWL) e visualmente representada na ferramenta Protege¹⁰. Na Figura 2, são mostrados todos os recursos e as ações em alto nível com nomes padronizados. Isso busca facilitar o tratamento de futuras aplicações desenvolvidas tendo como base a

⁶aws.amazon.com

⁷cloud.google.com

⁸www.openstack.org

⁹github.com/openstack

¹⁰protege.stanford.edu

Tabela 1: Exemplo de mapeamento na ontologia para o recurso Image

AWS	GCP	OS	Ontologia
Copy, Create, Import, Register	Create	Add, CopyFrom, Upload	Create
—	Get, GetFromFamily, GetIamPolicy	Download, Get	Read
Deregister, Register	Deprecate, SetIamPolicy, Update, UserReadOnly	Communitize, Modify, Publicize	Update
Deregister	Delete	Delete	Delete
—	List	—	List
—	SetLabels	—	Add
—	SetLabels	—	Remove

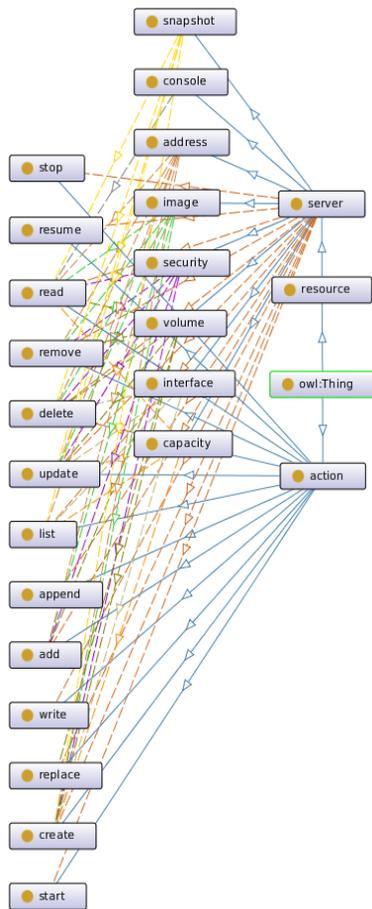


Figura 2: Primeira versão da ontologia: Conceitos, atributos e relacionamentos

ontologia, como Kumo que é um serviço especializado na realização de migrações de máquinas virtuais em ambientes multi-nuvem heterogêneos.

7 KUMO

Nesta seção, o serviço Kumo, sua API e drivers são detalhados. Kumo é a palavra na língua japonesa para nuvem e é representado

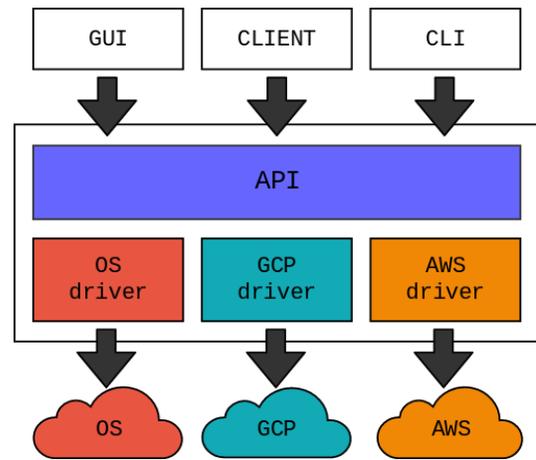


Figura 3: Arquitetura do Kumo conductor

pele caractere do alfabeto Kanji 雲. Kumo é responsável por implementar e expor uma API web que permitirá aos usuários registrar suas VMs e também acionar outras ações, como migrar. Para isso, Kumo possui o conceito de drivers, que são os especialistas em realizar operações necessárias à migração.

API destina-se a expor a execução remota de código e é necessário para resolver o problema de heterogeneidade através de sua interface unificada, em que os usuários solicitarão migrações para as nuvens integradas ao Kumo, de acordo com as definições da ontologia;

Drivers são as implementações de código que executam as ações necessárias nas nuvens de origem e destino (por exemplo, autenticação). É a maneira de tornar o Kumo capaz de interoperar com as múltiplas nuvens;

Conductor é a aplicação responsável por iniciar a API e executar as ações definidas e implementadas pelos drivers, bem como realizar o controle de ciclo de vida das ações definidas pelo serviço.

Tendo apresentado os principais componentes que compõem o Kumo, é interessante entender o fluxo onde um usuário solicita uma migração, até o estágio em que a VM é provisionada e entra em execução na nuvem destino. Na Figura 4, é mostrado um processo de alto nível de migração de uma máquina virtual.

Na primeira ação da Figura 4, (1) solicita a migração de uma VM da nuvem de origem para uma nuvem de destino por meio da API

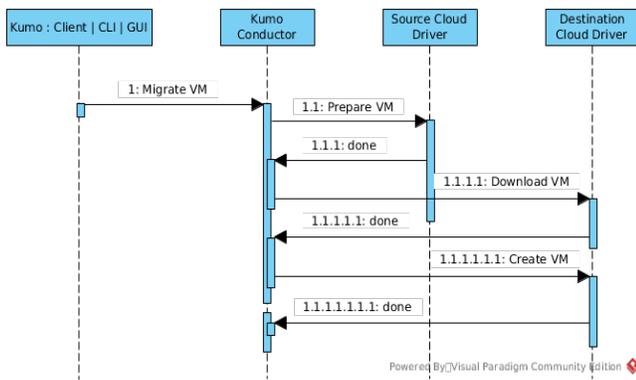


Figura 4: Diagrama de sequência da migração usando Kumo

do Kumo executada pelo conductor. Em (1.1), o processo de preparação de VM na nuvem de origem é iniciado pelo driver da nuvem de origem. A fase de preparação pode incluir o desligamento da VM; criação de um *snapshot*; buscar configurações de rede e armazenamento; obter metadados. Após essas etapas, o Kumo conductor recebe todas as informações necessárias sobre essa VM da nuvem de origem (1.1.1). Caso nenhum erro ocorra, o próximo passo é acionado (1.1.1.1). Nessa fase, o driver Kumo da nuvem de destino executará o registro do arquivo da VM e também replicará todas as configurações que foram coletadas da nuvem de origem para a de destino. Como o conductor sabe que todas as configurações foram feitas com sucesso (1.1.1.1.1), ele solicita a criação da VM na nuvem de destino (1.1.1.1.1.1), garantindo também todas as configurações que realizamos na VM e com isso a migração termina (1.1.1.1.1.1.1).

7.1 Ferramentas Kumo

Nesta seção, as ferramentas que permitem a operação do Kumo são detalhadas. Essas ferramentas possibilitam a conexão com o Kumo, e o desenvolvimento de uma aplicação para execução no serviço, permitindo uma melhor interação dos operadores com o serviço.

7.1.1 Kumo Client. Os usuários do Kumo podem precisar criar seus próprios aplicativos e *scripts* para automatizar suas migrações e cargas de trabalho. Por causa disso, o cliente Kumo será a maneira mais fácil de acessar, gerenciar e migrar os recursos por meio de código.

7.1.2 Kumo CLI. Para ter uma melhor experiência de usuário, os operadores que precisam gerenciar manualmente grandes quantidades de recursos, seja através de *scripts* de automação ou por meio do terminal do sistema operacional.

7.1.3 Kumo GUI. Opcionalmente, será possível operar o serviço Kumo para executar migrações usando um navegador web através da sua interface gráfica (GUI). Após a integração desse componente ao ecossistema Kumo, os usuários poderão gerenciar com mais rapidez e facilidade seus recursos em várias nuvens, mesmo usuários sem muita experiência em operações.

8 CONCLUSÕES

Atualmente, a primeira versão de uma ontologia para ambientes multi-nuvens heterogêneos com foco na portabilidade de máquinas virtuais foi desenvolvida. A abordagem utilizada para o desenvolvimento de tal ontologia foi *bottom-up*, uma vez que foram analisadas políticas de autorização de três nuvens heterogêneas para que fossem extraídos os recursos e ações comuns a elas, chegando assim ao que se chama de *core* da ontologia.

Também foi detalhada uma proposta de arquitetura para o serviço que implementará os conceitos definidos pela ontologia. Este serviço chamado Kumo, será uma outra contribuição da pesquisa e também será responsável pela portabilidade de máquinas virtuais entre nuvens heterogêneas. Neste trabalho, é definida a arquitetura do serviço Kumo e seus componentes que serão implementados durante a execução da pesquisa.

Devido a sua arquitetura flexível, outros conceitos poderão posteriormente ser integrados ao serviço Kumo, como a possibilidade de gerenciamento de infraestrutura de maneira unificada ou ainda a adição de capacidades de migração de outros componentes de infraestrutura, como contêineres. Para isso, é necessária previamente a evolução da ontologia para mapear os novos recursos, bem como suas ações.

9 AGRADECIMENTOS

Este trabalho foi possível pela bolsa de mestrado fornecida pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Programa PROEX, Edital 0487 e Processo 1741270.

REFERÊNCIAS

- [1] Nick Bassiliades, Moisis Symeonidis, Panagiotis Gouvas, Efstratios Kontopoulos, Georgios Meditskos, and Ioannis Vlahavas. 2018. PaaS semantic model: An ontology for a platform-as-a-service semantically interoperable marketplace. *Data & Knowledge Engineering* 113 (2018), 81–115. <https://doi.org/10.1016/j.datak.2017.11.001>
- [2] B. Chandrasekaran, J. R. Josephson, and V. R. Benjamins. 1999. What are ontologies, and why do we need them? *IEEE Intelligent Systems and their Applications* 14, 1 (Jan 1999), 20–26. <https://doi.org/10.1109/5254.747902>
- [3] Divyaa Manimaran Elango, Frank Fowley, and Claus Pahl. 2018. An Ontology-Based Architecture for an Adaptable Cloud Storage Broker. In *Advances in Service-Oriented and Cloud Computing*, Zoltán Ádám Mann and Volker Stolz (Eds.). Springer International Publishing, Cham, 86–101.
- [4] Object Management Group. 2017. *Interoperability and Portability for Cloud Computing: A Guide* (2.0 ed.). Object Management Group, <https://www.omg.org/cloud/deliverables/CSCC-Interoperability-and-Portability-for-Cloud-Computing-A-Guide.pdf>
- [5] Kate Keahey, Pierre Riteau, and Nicholas P. Timkovich. 2017. LambdaLink: An Operation Management Platform for Multi-Cloud Environments. In *Proceedings of the 10th International Conference on Utility and Cloud Computing (UCC '17)*. ACM, New York, NY, USA, 39–46. <https://doi.org/10.1145/3147213.3147224>
- [6] OpenStack. 2018. Open source software for create private and public clouds. (2018). <https://www.openstack.org>
- [7] R. Pellegrini, P. Rottmann, and G. Strieder. 2017. Preventing vendor lock-ins via an interoperable multi-cloud deployment approach. In *2017 12th International Conference for Internet Technology and Secured Transactions (ICITST)*. 382–387. <https://doi.org/10.23919/ICITST.2017.8356428>
- [8] Statista. 2018. Public cloud service usage worldwide 2018. (2018). <https://www.statista.com/statistics/511467/worldwide-survey-public-coud-services-running-application>
- [9] Adel Nadjaran Toosi, Rodrigo N. Calheiros, and Rajkumar Buyya. 2014. Interconnected Cloud Computing Environments: Challenges, Taxonomy, and Survey. *ACM Comput. Surv.* 47, 1, Article 7 (May 2014), 47 pages. <https://doi.org/10.1145/2593512>

A Framework for Automatic Topic Segmentation in Video Lectures

Eduardo R. Soares

Post Graduate Program in Computer Science
Federal University of Juiz de Fora
Juiz de Fora, Minas Gerais
eduardosoares@ice.ufjf.br

Eduardo Barrère

Post Graduate Program in Computer Science
Federal University of Juiz de Fora
Juiz de Fora, Minas Gerais
eduardo.barrere@ice.ufjf.br

ABSTRACT

Nowadays, video lectures are a very popular way to transmit knowledge, and because of that, there are many repositories with a large catalog of those videos on web. Despite all benefits that this high availability of video lectures brings, some problems also emerge from this scenario. One of these problems is that, it is very difficult to find relevant content associated with those videos. Many times, students must watch the entire video lecture to find the point of interest and, sometimes, these points are not found. For that reason, the proposal of this master's project is to investigate and propose a novel framework based on early fusion of low and high-level audio features enriched with external knowledge from open databases for automatic topic segmentation in video lectures. We have performed preliminary experiments in two sets of video lectures using the current state of our work. The obtained results were very satisfactory, which evidences the potential of our proposal.

KEYWORDS

Topic segmentation; Video lectures; Automatic Speech Recognition, Semantic annotation, Knowledge base, Content processing, Natural Language Processing

1 INTRODUCTION

Although videos have always been considered a powerful source to transmit knowledge, nowadays its importance is even more present in everyday life. That is mainly due to the fact that advances in multimedia and communications provided the means to creation of interactive and robust online educational systems where video lectures are made available [24]. The academia, for example, has widely embraced e-learning model. Many universities offer distance learning courses aiming to reach students who do not have access to the campus. Just as companies which want to provide training to its employees and ensure that they learn new abilities without taking too long for that.

Unluckily, the popularization of digital video has brought with it the difficulty of users finding relevant content, on video repositories, according to their interests. This is due to the fact that there is an overload of information in video format on the web [9]. For this reason, one of the biggest challenges that information retrieval

researchers face today is allowing users to access relevant information for their searches in the midst of so much content available. When we talk specifically about video lectures, it is very common that students watch the video to learn about just a few specific topics. But, typically, to quickly access this information is not an easy thing to do. Frequently, students spend a lot of time to localize interest points in a video. This occurs, mainly, because of the unstructured and linear nature of video that does not provide navigability through contents that is ideal for learning [11].

Topic segmentation is the most common preprocessing to allow navigability through video lecture content. And it is consensual that this is able to turn information retrieval more agile. In addition, a topic-segmented video lecture can enhance learning across distance learning platforms by providing students with structured video content so they can navigate instantly from one topic to another whenever they choose, at their own learning pace [23, 24].

Despite all the advantages of segmenting video lectures into topics, that is not a trivial task. Human segmentation is very accurate, but the time required to accomplish that task manually is almost impractical, especially in large repositories of video lectures [11]. That is why many methods were proposed in the literature to automatically extract topic structure in video lectures through performing analysis in their features. Although these methods can achieve acceptable results, they ignore the fact that there is background information that cannot be obtained just with features of the video itself and that could be used to improve topic segmentation. For example, there are some external open knowledge bases like DBpedia [2] that can be used to enrich video features with this objective. Following this reasoning, we propose the use of low and high-level audio track features supported by external knowledge through an early fusion to obtain a topic segmentation for video lectures. Our main hypothesis is that, as our proposal uses different levels of audio track features added with external knowledge, it can obtain different relevant information from video lecture that combined contribute to a more complete understanding of the addressed topics in the video. And, that understanding may be used to obtain an accurate segmentation. Furthermore, when available, textual features from slides, textbooks, and metadata can also be used by our framework to improve topic segmentation.

In this work, we consider the automatic topic segmentation task as: given a video lecture V as input, determine the start and end time of each existing topic T_i in V . Where, we define a topic as a unit of a video lecture that is composed of sequential chunks that cover a same subject.

This paper is organized as follows. In section 2 we talk about the related researches of literature that deal with automatic topic

In: XVIII Workshop de Teses de Dissertações (WTD 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pôsteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.
© 2018 SBC – Sociedade Brasileira de Computação.
ISBN 978-85-7669-435-9.

segmentation problem for video lectures. Section 3 describes our proposal and the ideas behind it. Section 4 describes the current state of our work and the preliminary results. Finally, in section 5 our conclusions are shown.

2 RELATED RESEARCH

Due to the relevance of the subject of this present work to the multimedia and information retrieval area, many different approaches to automatically extract topic structures in video lectures have been proposed over the years. These approaches, generally, may make use of different modalities of video information (eg. audio, video, text), in different semantic levels (eg. high, mid and low-level).

For example, in [4], the authors proposed a method for automatically obtaining video lecture summaries through topic segmentation. Their method is based on highlighting important sentences that were spoken in the video. To do this, they extracted low-level acoustic characteristics and used them to assign an “importance factor” to each of these sentences. The characteristics used by them had already proved to be good at detecting a change of subject in spoken discourses by identifying points where the speaker emphasized [1, 5]. Some of them are pitch, volume, duration of syllable sounds, and pause rate. Just like in [4], Togashi et al. [19] also made use of those same acoustic features, but combined with higher level linguistic features like cue words and phrases, word repetition, terms frequencies, and sentence locations. As results, the authors reported that the combination of those low and high-level features performed better than using them separately.

Speaking of methods that use high-level language structures to detect topic boundaries in video lectures, Lin et al. [11] proposed a whole linguistics-based one. In their approach, there are two sets of features which they aim to extract: content-based and discourse-based. The content-based set is composed of linguistic structures like noun phrases, verb classes, word stems, and others. Unlike it, the discourse-based set is composed of pronouns and cue phrases. According to the authors, the content-based structures are more related to the lexical and syntactical meaning of the body of content, while the discourse-based set has more to do with the neighborhood of hypothetical topic boundaries. After that feature extraction, a vector space is built using the weights of each feature in fixed size windows of transcript sentences. Then the similarity between each neighbor window is calculated and, the final topic segmentation is obtained using a similarity criterion.

When we work with video lectures, it is necessary to consider the existence of some specific characteristics that may be used to improve topic segmentation. One of them is that, frequently, video lectures follow the content of a textbook. Thus, Yamamoto et al. [22] presented a method for topic segmentation in video lectures through the association of audio transcription coming from automatic speech recognition (ASR) with topics keywords from the textbook summary. For this, they create a vector space where there are vectors that represent transcribed sentences and that represent textbook topics. Then the similarities between sentences vectors and textbook topics vectors are calculated. And, each sentence is associated with a single textbook topic using those similarities. Although, the assumption that the textbooks will always be available is a great disadvantage.

Another important characteristic of video lectures is the consensus that they, generally, do not present any significant visual changes like it is notorious in other kinds of video (e.g movies, news, cartoons, etc) [11]. Yet, there are some visual features in this kind of video that can be useful in automatic topic segmentation. That is why researches were carried out in order to explore those features, either individually or by combining with other sources of information. Like in [10], where the authors proposed a method based on image processing techniques to automatically extract handwriting from the blackboard and, thus, to identify the cutting points of a video lecture to obtain a topic segmentation. In [15], a combination of visual and text features was proposed. In this approach, three sources of information are used: video, slides, and subtitles. For each one, the method searches for transitions cues that indicate a topic change. Lastly, the transitions points obtained from the previous step are merged to form the final topic segmentation. As results, the authors reported that the combination of those transitions points improved significantly some evaluation ratios.

In [17], the focus is to automatically summarize lectures slides. The motivation behind that work is that it is very common for students to be asked to prepare for classes before they happen. The authors presented a method that can improve students preview through the use of visual and textual resources. To evaluate their approach in a real application, the authors conducted researches with more than 300 students. Their findings show that the use of summarized lectures slides by their method did not impact negatively on students performance and were capable of reducing their preview time. Slides can be a rich source of information in a lecture video. Furini et al. [7] published a recent work where they proposed the use of low-level audiovisual features combined with Optical Character Recognition (OCR) on slides to obtain a topic-based playlist, and thus to improve the information search on video lectures. Points of topic change are detected by low-level audiovisual features, and OCR is performed on slides to extract content information and allow keyword searches.

Still talking about multimodal approaches, Kishi and Goularte [9] proposed a method for automatic video scene segmentation where the features of different information channels are combined by computing their co-occurrences in shots before the segmentation step. This type of approach is called early fusion. Otherwise, if the approach takes individual decisions for each information channel and, after, those decisions are combined, it is called a late fusion approach. Although the Kishi and Goularte [9] method was proposed for automatic video scene segmentation, it provides a generic way to combine multiples sources of information that can be also used to segment video lectures into topics.

So far, we have only mentioned works that proposed methods which extract all information from video lecture itself. But, as we defend, there is background information that cannot be obtained just from the video. In this sense, there are researches of literature that have explored the use of external knowledge bases on topic segmentation task. In the work of Lin et al. [12], was presented an approach for automatic topic segmentation in video lectures that makes use of low-level audiovisual features, speech transcript, plus information from two types of lexical knowledge bases. The first one is a base of general words that are organized in synonyms

sets that are connected by semantic relationships, while the second is analogous to the first but specialized for lecture domain. More recently, a late fusion approach, that uses a knowledge base, was proposed by Shah et al. [16]. Their approach computes topic boundaries using subtitles, visual features, and Wikipedia articles, in a separate way. Then those topic boundaries are combined to generate the final set of topic boundaries. To find topic boundaries using subtitles and visual features, they used already cited methods from [11, 15]. To get topic boundaries using Wikipedia articles, they proposed a novel method that consists of segmenting into blocks the Wikipedia article that has the same subject as the video. And then, the method finds blocks of words from subtitles that most closely matches with Wikipedia article blocks to be the topic boundaries.

In this section, we briefly presented the state-of-the-art researches in automatic topic segmentation for video lectures. It was important to understand the tendencies of this area and situate the readers about the inspirations that we had when proposing our framework. In this paper, we propose a framework that makes use of early fusion of audio track low and high-level features combined with external knowledge. The advantage of our proposal is that it does not make assumptions about the existence of slides, textbooks or metadata associated with the video lecture, but it is capable to incorporate those information sources, when available, to improve the automatic topic segmentation.

3 PROPOSAL OF MASTER’S PROJECT

What we propose as master’s project is a novel framework that is able to extract, enrich semantically and combine features from different sources, at different semantic levels (low and high), with the objective of segmenting video lectures into topics. Some of those features will not be essential for the operation of the framework, and will only be considered when their source is available for a specific video lecture. The semantic enrichment will be done through knowledge bases from where video lecture concepts and their relationships will be retrieved. In the end, an algorithm will take the segmentation decisions based on all extracted information. Figure 1 illustrates the proposed framework.

The main idea for obtaining this framework is to conduct an extensive study about techniques and methods of extraction, pre-processing, semantic enrichment and combination of video lecture features. Furthermore, the algorithms that use those features to segment the video lecture into topics will also be studied and analyzed. In the end, we expect to have a generic framework where it is possible to combine features from different sources related to a video lecture, when available, to obtain an accurate topic segmentation. The only mandatory source to the framework will be the audio track from which low and high-level features will be extracted since most of the information in a video lecture is in the teacher’s speech.

3.1 Sources of information

Video lectures may have multiple sources of information, and thus, it is needed that the framework can handle that. The audio track will be the main source of information. Since the content of teacher’s speech has a great importance on identifying the subjects of a video lecture. Thus, the audio track is expected to always be available.

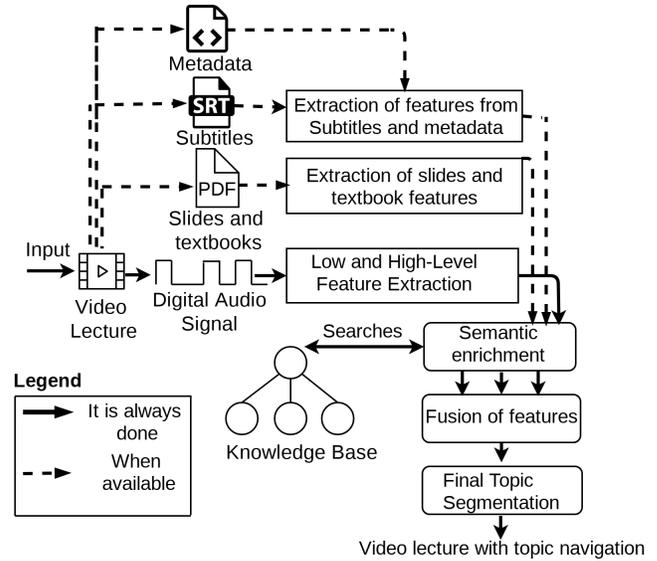


Figure 1: Proposed Framework

Other sources of information that will be considered by our framework, when available, are subtitles, slides, textbooks, and metadata such as LOM (Learning Object Metadata) [8].

Different from other literature approaches, our proposal is only dependent on the audio track. The other sources of information are not essential to framework operation but can be used when available to improve topic segmentation by bringing more information about video lecture’s content.

3.2 Extraction of features

The extraction of features from information sources is a very important process of our framework. In this step, the framework extracts features that will be semantically enrichment and used by the topic segmentation algorithm.

For audio track, low and high-level features can be extracted. Some of the possible low-level features that will be considered by the framework are silence detection, fundamental frequencies (f_0) and voice power density estimations. As stated in the previous section, these features have proven to be efficient in detecting moments of emphasis on speeches. And, these moments provide a good signal that there has been a change in subject. Since speakers tend to emphasize when they start a new topic. Another important low-level feature that we intend to analyze its contribution in topic segmentation task is the Mel-frequency cepstral coefficients (MFCC) [14]. Its use is widely diffused in the extraction of acoustic characteristics for automatic speech recognition. Therefore, MFCC is capable of giving a good acoustic representation along the audio track, which can be useful in topic segmentation task.

About the high-level features from the audio track, we intend to extract them in text format, through automatic speech recognition (ASR). The audio transcription is a powerful source of information because it makes possible to obtain the audio content in a format that allows the comparison between the contents of parts of the video, which can be used to identify a change of topics in the

video lecture. For the other sources of information, only high-level features will be extracted, also in text format. The textual content of metadata and subtitles can be extracted by parsing plain text files. For textbooks and slides, when the text content is rendered in PDF file, specific parsers can be used. But if the content is in Raster Matrix Format (RMF), both externally and within the video lecture, OCR (Optical Character Recognition) can be used for extraction [7].

All these textual content we could extract from video lecture may be submitted to Natural Language Processing (NLP) techniques like tokenization, stop-words removal, stemming, POS-tagging and others [3]. Furthermore, there are also well-known techniques to represent textual contents in a way that is possible to compute their similarities, like Bag-of-Words (BoW) and N-gram model [21].

3.3 Semantic enrichment

The semantic enrichment of features is a framework process where concepts found in feature extraction step will be searched in a knowledge base and, then, enriched with their relationships with the objective of adding context to them. The justification for this is that, without context, some concepts may seem to be unrelated, but when we analyze their relationships in a knowledge base we may find the opposite. And, that finding can improve topic segmentation decisions. For example, suppose that the concept “TCP” (Transmission Control Protocol) is extracted, through ASR, at the 30 seconds of a computer networks lecture. Later, at the 60 seconds of the same video lecture, the concept “UDP” (User Datagram Protocol) is extracted. Without any context, the topic segmentation algorithm may take the decision to separate those two video parts on different topics. But if we add some context to those concepts, we are able to discover that both TCP and UDP are Transport Layer Protocols and, therefore, are related. Thus, the segmentation algorithm may think it is better to change its previous decision and do not separate the video parts where those concepts were extracted.

To extract concepts from the features of the video lectures and enrich them, we intend to use the semantic annotation methods allied to searches in a knowledge base. To perform this, open and free knowledge bases like DBpedia[2] are available and can be used.

3.4 Topic segmentation

Choose an appropriate topic segmentation algorithm is an important issue to be solved in this master’s project. There are two main classes of algorithms that can be used to perform that task: clustering and classification algorithms. The clustering algorithms will focus on try to find similarity and dissimilarity between parts of the video lecture, based on extracted features, to generate the video lecture topics. Well-known clustering algorithms that can be used are: *K-means*, *K-medoids*, *DBSCAN*, *Spectral Clustering*, and others [6, 20]. Instead, classification algorithms will use the extracted features to identify the existence or otherwise of a subject change in given segments of the video lectures. Classification algorithms that can be explored in this master’s project include: *Decision Trees*, *Support Vector Machines*, *Neural Networks*, *Naive Bayes classifiers*, and others. In addition, because the performance of the algorithms can be influenced by the features used, as well as by the way we

combine them, the evaluation of the techniques of early fusion of features is also part of our plans.

4 CURRENT STATE OF THE WORK

In this section, we will briefly discuss the current state of our work. First, we will present the current stages of processing performed by the framework. Then we will present the preliminary experiments that were performed and discuss the results obtained. A more detailed explanation can be found at [18].

4.1 Current framework processing

The current processing steps of our framework can be seen as a pipeline. In other words, the output of each stage of processing is the input of the next. The main idea of this approach is that, as the media flows through the pipeline, it is processed and transformed, allowing different semantic levels of information to be extracted from it. An overview of the proposed approach can be seen in Figure 2. From now on, we will use this figure as a reference to explain how the video lecture processing occurs.

To start the process, a video lecture is given as input and its audio track is extracted (i). Next, the extracted audio track is divided into chunks. That division is made so that the generated chunks do not contain silence, that is, after this stage, we have a sequence of fully voiced audio chunks (ii). After that, for each of these chunks, we extract a feature vector that relates the fundamental frequencies (f_0) and power spectral density (PSD) [13] that occur over the chunk. Then, we use those vectors to compute and store an affinity matrix M_f that indicates the similarity between every pair of chunks according to those features (iii). In the next step (iv), those chunks are transcribed by an automatic speech recognition system. Thus, we obtain what was said in each chunk in text format. So in the fifth stage (v), we build a vector space with those transcriptions. In this space, each chunk s_i , where $i \in [0, 1, 2, \dots, N - 1]$ and N is the number of chunks, is represented by a vector and each dimension of this vector gives the weight of a word w_j in s_i . Where $j \in [0, 1, 2, \dots, V - 1]$ and V is the vocabulary size. This vocabulary can be predefined, or it can be built on the words transcribed in the video lecture for memory savings purpose. After that, another affinity matrix M_t is computed and stored. But this time, by considering the transcription features. The last semantic level of information extraction is done in steps (vi) and (vii). In the sixth step, each text from audio transcription is submitted to a semantic annotation. So after that, for each audio chunk s_i , we have a set of annotated terms A_i . Then, in the seventh step, for each set of annotated terms A_i , and for each annotated term $a_m \in A_i$, where $m \in [0, 1, 2, \dots, G_i - 1]$ and G_i is the number of annotated terms in A_i , we search in a knowledge base for terms that have the same meaning as a_m . Next, we take those terms, including a_m , and we get their categories in the same base. After that, we have for each term $a_m \in A_i$ its synonyms and categories. So in the step (viii) is built a vector space and calculated an affinity matrix M_a , just like in step (v), but this time considering the features that were extracted in the steps (vi) and (vii). In the step (ix), we linearly combine all affinity matrices that were obtained in the previous steps to get a final affinity matrix H , used in the step (x) by the spectral clustering algorithm to generate the video lecture topic boundaries. These

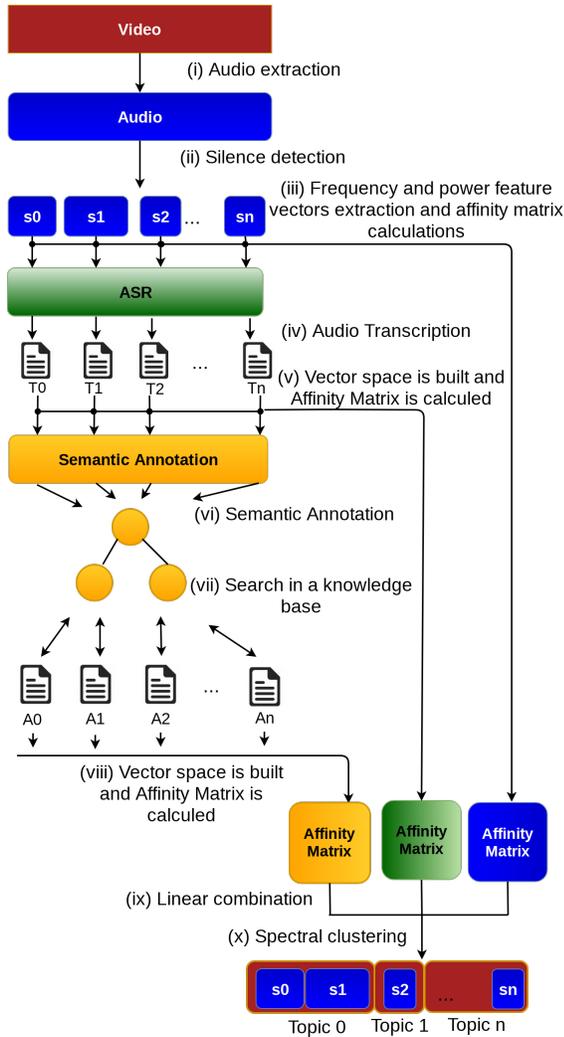


Figure 2: Overview of the current stage of the framework [18]

boundaries are obtained by clustering audio chunks according to H and, then, by using timestamps to map clusters of audio chunks into video lecture topics.

4.2 Preliminary Experiments

We conducted preliminary experiments with the objective of evaluating the impacts of incremental addition of higher-level features for topic segmentation. Thus, we separated the features of video lecture’s audio track in three categories, from the lowest level to the higher: F , T and A . Where F are the frequency and power features, T represents the audio transcription features and A the features extracted through semantic annotation. Then, we consider that we have a sequence C of 3 sets formed by the combination of those categories in a way that, the set $i + 1$ of C has the elements of the set i plus one element of the next category. That is $C = \{\{F\}, \{F, T\}, \{F, T, A\}\}$. Then, for each video lecture v_j from

the data set, we run the framework processing 10 times considering each set of features. So, for each set of features, we calculate the average of evaluation ratios obtained for that video lecture through the 10 executions. With this experimentation, we want to show that combining lower level features with higher ones increases the quality of topic segmentation. The need to execute our method 10 times for each set of features in each video lecture is that the used clustering algorithm is based on K-means, which is a random start algorithm. So by running Q times and taking the average, we obtain a more significant evaluation of our method.

4.2.1 Evaluation data set. Our evaluation dataset is composed of 44 video lectures in Brazilian Portuguese, where 34 of them were extracted from the Videoaula@RNP repository¹. The video lectures from this repository already have a topic segmentation which we used as the ground-truth for comparison. The other 10 video lectures were extracted from YouTube, and they did not have a previous segmentation. We had to make our own manual segmentation to use as ground-truth in this case. The decision to merge videos from those two sources into a dataset was based on the fact that the video lectures from Videoaula@RNP were made in a more traditional way by following a well-defined script, unlike the YouTube videos selected. The YouTube video lectures were made in a freer style and have a shorter duration than traditional video lectures. In this way, we also want to evaluate the impact of those differences on our method’s performance.

Since in our research we did not find any evaluation data set in Portuguese for automatic topic segmentation task, we make our dataset publicly available on Google Drive². For the final project of master’s degree, we also intend to evaluate our framework in a data set of video lectures in English.

4.2.2 Evaluation ratios. To evaluate our method, we compare for each video lecture the automatic segmentation with the ground-truth. For these comparisons, we have chosen the mean of precision, recall e F-measure across the ground-truth topics. These are typical ratios used in literature to evaluate the results of information retrieval and multimedia processing tasks. Along the development of the master’s project, other metrics that can be good for evaluating the framework’s performance in the task of automatic topic segmentation, such as Overflow and Coverage [9], can also be used.

4.3 Preliminary Results

By doing the experiments that were previously described, we got the overall results which are shown in the Table 1 and 2. While Table 1 shows the results obtained in the data set of video lectures from Videoaula@RNP, Table 2 presents the obtained results on the YouTube videos. As can be seen, in overall, the incremental addition of feature levels has improved the automatic topic segmentation in 5%, if we compare the first set of features $\{F\}$ with the last one $\{F, T, A\}$. This behavior was noticed in the two data sets, which is a good sign that by considering different levels of information we can get a more accurate topic segmentation.

Also, in this preliminary experimentation, it has to be noted that the overall obtained results in the two data sets were significantly

¹<http://www.videoaula.rnp.br/portal/home>

²<https://goo.gl/UFG88k>

Table 1: Method performance on Videoaula@RNP videos

Set of Features	{F}			{F, T}			{F, T, A}		
Lecture	Prec.	Rec.	Fm.	Prec.	Rec.	Fm.	Prec.	Rec.	Fm.
Overall	0,59	0,61	0,57	0,63	0,64	0,60	0,64	0,65	0,62

Table 2: Method performance on YouTube video lectures

Set of Features	{F}			{F, T}			{F, T, A}		
Lecture	Prec.	Rec.	Fm.	Prec.	Rec.	Fm.	Prec.	Rec.	Fm.
Overall	0,85	0,80	0,80	0,89	0,83	0,83	0,90	0,85	0,85

distinct. According to our investigations, the reason for that is because the duration of the video lectures and their number of ground-truth topics impact directly on the difficulty of segmenting them into topics. This occurs because of a characteristic of the problem, in which errors in the formation of a topic are propagated to the following topic, then the final propagated error is directly proportional to the number of topics in the video lecture. That error propagation occurs because of when the method wrongly clusters an audio chunk s_j in a topic t_i , the precision of t_i is affected because of that error, just like the recall of t_{i+1} , where s_j should belong to. Thus, the average F-measure of an obtained topic segmentation also decreases. And, in our data set, the video lectures from YouTube has, on average, fewer topics than those from Videoaula@RNP. Another conclusion drawn from our analysis in the evaluation data set is that the videos extracted from YouTube present, in general, a much higher audio quality in relation to those from Videoaula@RNP. And, as all stages of feature extraction of our framework, that were implemented so far, are sensitive to noise, it is natural that a low quality of recording leads to more errors than a high one.

5 CONCLUSION

In this paper, we presented a master's project proposal which consists of a framework for automatic topic segmentation for video lectures. In our proposal, different levels of features from the audio track of a video lecture can be extracted and combined with features from other sources related to it. Like slides, textbooks, and metadata. We conducted preliminary experiments over the current stage of work that served to evidence for the potential of combining audio features of different levels and semantically enriching them through a knowledge base.

The next steps of this master's work, until the limit of March of 2020, consist of: constant updating of the bibliographic review; study and implementation of feature extraction techniques to explore the other sources of information; analysis of pre-processing and noise removal techniques; implementation and evaluation of techniques for early fusion of features and final topic segmentation; and lastly, the writing and defense of dissertation, and the publication of the results in papers.

REFERENCES

- [1] Barry Arons. 1994. Pitch-based emphasis detection for segmenting speech recordings. In *Third International Conference on Spoken Language Processing*.
- [2] Sören Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives. 2007. Dbpedia: A nucleus for a web of open data. In *The semantic web*. Springer, 722–735.
- [3] Steven Bird and Edward Loper. 2004. NLTK: the natural language toolkit. In *Proceedings of the ACL 2004 on Interactive poster and demonstration sessions*. Association for Computational Linguistics, 31.
- [4] Xiaoyin Che, Sheng Luo, Haojin Yang, and Christoph Meinel. 2016. Sentence-Level Automatic Lecture Highlighting Based on Acoustic Analysis. In *Computer and Information Technology (CIT), 2016 IEEE International Conference on*. IEEE, 328–334.
- [5] Francine R Chen and Margaret Withgott. 1992. The use of emphasis to automatically summarize a spoken discourse. In *Acoustics, Speech, and Signal Processing, 1992. ICASSP-92., 1992 IEEE International Conference on*, Vol. 1. IEEE, 229–232.
- [6] Adil Fahad, Najlaa Alshatri, Zahir Tari, Abdullah Alamri, Ibrahim Khalil, Albert Y Zomaya, Sebt Foufou, and Abdelaziz Bouras. 2014. A survey of clustering algorithms for big data: Taxonomy and empirical analysis. *IEEE transactions on emerging topics in computing* 2, 3 (2014), 267–279.
- [7] Marco Furini, Silvia Mirri, and Manuela Montangelo. 2018. Topic-based playlist to improve video lecture accessibility. In *Consumer Communications & Networking Conference (CCNC), 2018 15th IEEE Annual*. IEEE, 1–5.
- [8] Wayne Hodgins and Erik Duval. 2002. Draft standard for learning technology-Learning Object Metadata-ISO/IEC 11404. *IEEE P1484. 12.2/D1* (2002).
- [9] Rodrigo Mitsuo Kishi and Rudinei Goularte. 2016. Video scene segmentation through an early fusion multimodal approach. *Anais do XXII Simpósio Brasileiro de Sistemas Multimídia e Web 2* (2016).
- [10] Greg C Lee, Fu-Hao Yeh, Ying-Ju Chen, and Tao-Ku Chang. 2017. Robust handwriting extraction and lecture video summarization. *Multimedia Tools and Applications* 76, 5 (2017), 7067–7085.
- [11] Ming Lin, Michael Chau, Jinwei Cao, and Jay F Nunamaker Jr. 2005. Automated video segmentation for lecture videos: A linguistics-based approach. *International Journal of Technology and Human Interaction (IJTHI)* 1, 2 (2005), 27–45.
- [12] Ming Lin, Christopher BR Diller, Nicole Forsgren, Yunchu Huang, and Jay F Nunamaker Jr. 2005. Segmenting lecture videos by topic: From manual to automated methods. *AMCIS 2005 Proceedings* (2005), 243.
- [13] Rainer Martin. 2001. Noise power spectral density estimation based on optimal smoothing and minimum statistics. *IEEE Transactions on speech and audio processing* 9, 5 (2001), 504–512.
- [14] Lindsalwa Muda, Mumtaz Begam, and Irraivan Elamvazuthi. 2010. Voice recognition algorithms using mel frequency cepstral coefficient (MFCC) and dynamic time warping (DTW) techniques. *arXiv preprint arXiv:1003.4083* (2010).
- [15] Rajiv Ratn Shah, Yi Yu, Anwar Dilawar Shaikh, Suhua Tang, and Roger Zimmermann. 2014. ATLAS: automatic temporal segmentation and annotation of lecture videos based on modelling transition time. In *Proceedings of the 22nd ACM international conference on Multimedia*. ACM, 209–212.
- [16] Rajiv Ratn Shah, Yi Yu, Anwar Dilawar Shaikh, and Roger Zimmermann. 2015. TRACE: Linguistic-Based Approach for Automatic Lecture Video Segmentation Leveraging Wikipedia Texts. In *Multimedia (ISM), 2015 IEEE International Symposium on*. IEEE, 217–220.
- [17] Atsushi Shimada, Fumiya Okubo, Chengjiu Yin, and Hiroaki Ogata. 2017. Automatic summarization of lecture slides for enhanced student preview-technical report and user study. *IEEE Transactions on Learning Technologies* (2017).
- [18] Eduardo R Soares and Eduardo Barrère. 2018. Automatic Topic Segmentation for Video Lectures Using Low and High-Level Audio Features. In *Proceedings of the 24th Brazilian Symposium on Multimedia and the Web*. ACM, 189–196.
- [19] Shingo Togashi, Masaru Yamaguchi, and Seiichi Nakagawa. 2006. Summarization of spoken lectures based on linguistic surface and prosodic information. In *Spoken Language Technology Workshop, 2006. IEEE*. IEEE, 34–37.
- [20] Ulrike Von Luxburg. 2007. A tutorial on spectral clustering. *Statistics and computing* 17, 4 (2007), 395–416.
- [21] Hanna M Wallach. 2006. Topic modeling: beyond bag-of-words. In *Proceedings of the 23rd international conference on Machine learning*. ACM, 977–984.
- [22] Natsuo Yamamoto, Jun Ogata, and Yasuo Ariki. 2003. Topic segmentation and retrieval system for lecture videos based on spontaneous speech recognition. In *Eighth European Conference on Speech Communication and Technology*.
- [23] Haojin Yang and Christoph Meinel. 2014. Content based lecture video retrieval using speech and video text information. *IEEE Transactions on Learning Technologies* 7, 2 (2014), 142–154.
- [24] Dongsong Zhang, Lina Zhou, Robert O Briggs, and Jay F Nunamaker Jr. 2006. Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information & management* 43, 1 (2006), 15–27.

Um Modelo de Predição para o Sucesso no Mercado Musical

Carlos V. S. Araujo
Universidade Federal do Amazonas
Instituto de Computação
Manaus, Amazonas 69080-900
vicente@icomp.ufam.edu.br

Rafael Giusti
Universidade Federal do Amazonas
Instituto de Computação
Manaus, Amazonas 69080-900
rgiusti@icomp.ufam.edu.br

Marco A. P. Cristo
Universidade Federal do Amazonas
Instituto de Computação
Manaus, Amazonas 69080-900
marco.cristo@icomp.ufam.edu.br

ABSTRACT

In this paper we present an ongoing research that aims to produce a model to predict success in the musical market. To reach this goal, it is necessary, initially, to identify influence factors in the market, which is the current focus of this research. In this branch, we identify that tweets have influence over the popularity of an album in Spotify. We also found out that Pop has more tendency to have the most popular song of a year, among other analyses. The next steps of this research are to list even more influence factors, generate the model using artificial neural networks and validate it with real-world cases.

KEYWORDS

Complex Network, Music Success, Artificial Neural Network

1 INTRODUÇÃO

A indústria do entretenimento (cinema, jogos, música e televisão) movimenta bilhões de dólares anualmente em todo o mundo. Segundo a Associação da Indústria de Gravação da América (*Recording Industry Association of America* – RIAA)¹, somente a indústria da música movimentou, em 2016, US\$ 7,7 bilhões nos Estados Unidos. Desse total, 75,5% são referentes às músicas digitais (*streaming* ou *download*).

Uma forma de maximizar os ganhos de artistas, gravadoras e selos musicais seria utilizando um modelo de predição de sucesso de um produto no mercado musical (álbuns, *singles*², DVDs etc.). Tal modelo poderia ser utilizado para se obter melhor resultado comercial, escolhendo a melhor época para seus lançamentos e uma melhor estratégia de marketing. A geração desse modelo e sua validação é o resultado final esperado ao fim dessa pesquisa em desenvolvimento. Para obtenção de tal resultado será necessário o uso de técnicas de mineração e fusão de dados, aplicadas a fatores de influência no mercado, tais como redes sociais, colaborações entre artistas e a popularidade do gênero musical.

As Redes Sociais Online (RSO), que vêm sendo adotadas por gravadoras para a divulgação de novos *singles* ou álbuns, são um desses fatores de impacto. Nessas RSO, os usuários têm a possibilidade de expor suas opiniões e percepções pessoais sobre os mais variados assuntos [16], sendo que estas podem vir a ser utilizadas como base

¹www.riaa.com/

²*Singles* são músicas lançadas individualmente em que se espera obter um bom resultado comercial.

In: XVIII Workshop de Teses de Dissertações (WTD 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pôsteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.

© 2018 SBC – Sociedade Brasileira de Computação.
ISBN 978-85-7669-435-9.

de influência para o comportamento de outros usuários [23], assim impactando na divulgação desses *singles* ou álbuns.

Trabalhos como os de Chen and Chellappa [5], Dewan and Ramaprasad [7] e Chen et al. [6] já mostram o impacto do conteúdo gerado por usuários em RSO no sucesso de um álbum musical.

Outro fator que vem sendo utilizado por artistas para obterem maiores vendas em suas músicas e álbuns são as colaborações, que ocorrem quando dois ou mais artistas diferentes participam de uma mesma música [3]. Trabalhos como os de Smith [22], Gleiser and Danon [12] e Uzzi [24] mapearam a colaboração de artistas em variados estilos musicais. Nesses trabalhos as colaborações foram mapeadas por meio de grafos, sendo os artistas representados como nós e suas colaborações por arestas.

A compreensão de como essas redes de relacionamentos são formadas e como evoluem ao longo do tempo possui elevado valor comercial. Esse conhecimento é fundamental para traçar eficazes estratégias de mercado para artistas consolidados e, principalmente, para novos artistas [3].

A popularidade do gênero musical no ano de lançamento de um álbum também é importante fator para sua popularidade. Mauch et al. [18] e Lopes [15] já estudaram como se dá o processo de popularização de um gênero, mostrando que revoluções são o fator principal para que um gênero passe a ter maior destaque. Mas, para que o interesse nele seja mantido é necessário que novos artistas de destaque nestes surjam a cada ano.

Utilizando esses fatores, além de outros ainda a serem estudados, é possível montar uma base de dados que pode ser usada para gerar um modelo de predição. Esse modelo se diferenciaria da literatura existente, pois em geral estes tratam apenas de fatores isolados, enquanto usaremos um englobamento desses elementos. Assim, buscando ser o mais próximo possível do valor real de sucesso alcançado pelo lançamento musical.

O restante deste trabalho está organizado da seguinte forma: na seção 2 são apresentados trabalhos relacionados ao desta pesquisa. Enquanto na seção 3 as contribuições esperadas a serem obtidas com esse estudo são mostradas. Na seção 4 os resultados obtidos e esperados são descritos. Por fim, na seção 5 está uma breve conclusão.

2 TRABALHOS RELACIONADOS

Os trabalhos de destaque nessa área de pesquisa voltaram-se a realizar predições utilizando-se de fatores isolados. Logo, nesta seção apresentamos alguns desses trabalhos e também aqueles que obtiveram fatores de influência como resultado, os quais podem vir a ser utilizados em nosso modelo.

Na área da análise de impacto de Redes Sociais Online no mercado musical, destacam-se os trabalhos de Dhar and Chang [9],

Dewan and Ramaprasad [8] e Chen et al. [6] que serão explanados a seguir.

Dhar and Chang [9] examinaram a viabilidade de usar conteúdos gerados por usuários, em redes sociais e blogs, na predição de vendas na indústria musical. Para tal, utilizaram uma amostra de 108 álbuns e acompanharam as mudanças sobre o que estava sendo falado a respeito deles durante as quatro semanas que antecederam e sucederam os seus lançamentos. Como resultado, descobriram que as vendas estão correlacionadas com o volume de postagens em blogs sobre um álbum.

Dewan and Ramaprasad [8] traçaram um paralelo entre as novas mídias (*downloads*), mídias antigas (CD's) e as vendas na indústria da música. Mais especificamente, estudaram a interação entre o que era comentando em blogs de distribuição gratuita de músicas, a quantidade de vezes que uma música era reproduzida na rádio e as vendas dessas músicas e de seus respectivos álbuns. Para análise, utilizaram a metodologia de autoregressão de vetores de painel (PVAR). Como resultados, demonstraram que o número de vezes que a música é tocada na rádio está diretamente relacionado a vendas futuras, tanto da música quanto do seu álbum.

No trabalho de Chen et al. [6], os autores empregaram o mesmo modelo utilizado por Dewan and Ramaprasad [8], a autoregressão de vetores de painel, com o intuito de investigar a relação entre divulgações em redes sociais e as vendas de música. Para isso, avaliaram os conteúdos automáticos e pessoais gerados na rede social MySpace³ e descobriram que a transmissão em mídias sociais tem um efeito significativamente positivo nas vendas. Mais especificamente, mostraram que esse efeito era ainda maior quando eram excluídas da análise as mensagens automatizadas.

Em relação à análise do processo de colaboração entre artistas na área musical, os trabalhos de Grandadam [13], Giaquinto et al. [11] e Budner and Grahl [4] se destacam.

Budner and Grahl [4] propuseram uma rede englobando os vários profissionais da indústria musical. Para montá-las utilizaram-se da lista “500 maiores álbuns de todos os tempos” da revista Rolling Stones⁴ e o livro “1001 álbuns que você deve ouvir antes de morrer”[10]. Nessa rede, cada profissional representava um nó e estes eram conectados se tivessem trabalhado em um mesmo disco. Ao analisar a rede gerada, os autores identificaram que as profissões mais importantes da indústria musical, em termos de conectividade, são os artistas principais e engenheiros de som.

Já Grandadam [13] e Giaquinto et al. [11] montaram redes para o gênero Jazz. O primeiro analisou o selo musical Blue Note, conectando na rede os artistas que participavam de uma mesma sessão de gravação. Então, utilizando-se da distância geodésica e do coeficiente de *clustering*, identificou que os artistas mais famosos colaboravam com novos artistas para que o *status* do selo se mantivesse. Enquanto o outro trabalho, de cunho mais generalista, apontou que para um artista obter sucesso nesse gênero é mais importante que este saiba tocar múltiplos instrumentos do que ter um valor de centralidade alto em sua rede de colaboração.

Na área de análise de popularidade de gêneros musicais, Mauch et al. [18] procuram identificar quais os principais fatores que implicam na popularização desses gêneros. Para isso, coletaram 86%

das músicas que apareceram no ranking “Billboard Hot 100”⁵ entre 1960 e 2010, onde as distribuíram em 13 diferentes grupos. Para tal, utilizaram as frequências de seus elementos, tanto instrumentais, quanto vocais. Por fim, os autores afirmam que a popularidade dos gêneros musicais é determinada por revoluções. Citam três destas como as de maior relevância: a invasão britânica de 1964, o movimento New Wave de 1983 e a ascensão do Rap em 1991, sendo essa última a de maior importância por ter mudado a estrutura musical americana.

Enquanto, Lopes [15] estuda a quantidade de novos e antigos músicos na lista de 100 *singles* mais populares entre os anos de 1969 a 1990. Apresentam que por ano ao menos 29,5% dos *singles* no Top 100 anual são de novatos, valor que chega até a 45%. Enquanto, aqueles com mais de oito anos de carreira representam no máximo 15,9%. Esses dados mostram que os novos artistas são fundamentais para a popularidade dos gêneros musicais.

Um outro aspecto a se considerar é o impacto das letras para o sucesso de uma música. Destaca-se aqui o trabalho de Marzagão [17], que analisa as letras de músicas brasileiras de oito diferentes gêneros musicais. Nesse trabalho, cada artista estudado recebe um número referente ao seu vocabulário, dado por uma razão entre a quantidade de palavras únicas utilizadas em suas composições e o valor total de músicas desse artista. O autor mostra que os artistas de Rap são os que tem maiores vocabulários em suas músicas, já os de Sertanejo tem o menor. Também chegou a duas conclusões: quanto menor o vocabulário do artista, maior tende a ser o seu sucesso; e, excluindo o Rap, quanto mais antigo é o artista, maior é o seu vocabulário.

Como foi apresentado, há uma gama de trabalhos que tratam de fatores de influência no mercado musical. Muitos desses trabalhos se atêm somente a mostrar que há esse impacto, não tentando prever qual seria o sucesso futuro de um álbum ou música a ser lançada. Já aqueles que realizam alguma forma de predição, utilizam somente fatores isolados. O diferencial da nossa proposta em relação aos trabalhos apresentados é o de englobar os vários fatores de influência, com o intuito de gerar um modelo de predição que seja o mais próximo possível do valor real de sucesso alcançado pelo produto musical.

3 CONTRIBUIÇÕES ESPERADAS

A grande contribuição esperada é voltada ao mercado musical, que utilizando-se do modelo poderá obter melhores rendimentos comerciais.

Entretanto, as contribuições acadêmicas serão advindas do processo requerido para a obtenção de tal resultado. Visto que serão necessárias técnicas de fusão e mineração de dados, computação social, aprendizagem de máquina e redes complexas.

Nesse quesito será primordial determinar a influência de parâmetros de influência no mercado musical que ainda não foram devidamente estudados, agregando conteúdo a literatura atual.

Por fim, o processo tomado para projetar o algoritmo que combine os resultados analíticos para determinação do modelo a ser proposto é significativo para a área de ciência da computação, mais

³<https://myspace.com/>

⁴<http://www.rollingstone.com/music/lists/500-greatest-albums-of-all-time-20120531>.

⁵<https://www.billboard.com/charts/hot-100>

precisamente a de ciência dos dados. Visto que será necessário aplicarmos redes neurais artificiais e encontrarmos aquela que venha a obter o melhor aprendizado dado o conjunto de fatores estudados.

4 RESULTADOS

Nessa seção apresentamos os resultados já obtidos e o que esperamos alcançar no decorrer desta pesquisa de mestrado, estando contidas nas subseções os resultados 4.1 e 4.2, respectivamente.

4.1 Resultados Obtidos

Na primeira etapa da pesquisa buscamos identificar fatores de influência no mercado musical. Essa etapa gerou três artigos publicados que serão explanados nesta seção.

Em Araujo et al. [2] identificamos uma correlação entre mensagens no Twitter⁶ (também conhecidas como *tweets*) escritas uma semana antes e após o lançamento de um álbum e seu respectivo valor em popularidade no Spotify⁷. Esse valor é um número entre 0 e 100 gerado pela própria plataforma com base no número de *streams* das músicas desse álbum.

Utilizamos a seguinte metodologia: inicialmente coletamos as mensagens utilizando a API da própria plataforma. Após, realizamos um pré-processamento, onde deixamos somente as mensagens únicas. Então, realizamos a detecção da polaridade dessas mensagens. Finalmente, analisamos se havia correlação entre os dados coletados e geramos dois modelos de predição. Tal processo está ilustrado na Figura 1.

Para mostrar que havia correlação entre as mensagens e o valor em popularidade no Spotify dos álbuns, aplicamos os coeficientes de correlação de Pearson, Spearman e de distância. Observamos que os dados apresentavam correlação linear positiva quando considerávamos o total de *tweets* ou somente aqueles com polaridade positiva. Valores ratificados com o teste-t que apresentou confiança acima de 95% nos resultados obtidos.

Utilizando tais dados, geramos dois modelos de predição: um utilizando a quantidade total de *tweets* e outros somente aqueles com polaridade positiva. Esses modelos foram avaliados considerando o coeficiente de determinação, a variância residual e o erro quadrático médio. Ao analisar esses três fatores, identificamos que o modelo das mensagens positivas apresentou melhor resultado. Ele é dado por:

$$P_{pos}(p) = 56,700 + \left(6,389 \times 10^{-4}\right)p, \quad (1)$$

onde $P_{pos}(\cdot)$ é a previsão da popularidade no Spotify, considerando a quantidade p de tweets positivos.

Em Araujo and Nakamura [1] utilizando os *rankings* “Year-End Hot 100 Songs”⁸ da Billboard entre os anos de 1969 a 2017 buscamos identificar fatores de influência sobre a popularidade dos gêneros musicais. Essas listas apresentam as 100 músicas mais famosas de cada ano nos Estados Unidos.

Realizada a coleta das músicas das listas, foi necessário identificar o gênero de cada uma delas. Para isso geramos um algoritmo que

realiza consultas à Wikipedia⁹ extraíndo essa informação. Ao montar representações contendo o número de músicas de cada gênero por ano, observamos os períodos nos quais havia uma repentina mudança nesses valores. Ao identificá-los, analisando informações sobre o mundo da música na época, conseguíamos extrair o motivo de tal acontecimento.

Assim, identificamos que o principal fator de influência sobre a popularidade de um gênero musical é o surgimento de um cantor, banda ou um movimento de destaque. Entretanto, em caso de músicos faz-se necessário que outros artistas também consigam obter notoriedade em anos próximos, fazendo com que o gênero consiga se manter em destaque. Caso isso não ocorra, possivelmente haverá um aumento repentino de popularidade, mas que não consegue se manter, dificilmente alcançando mais que cinco anos de notabilidade.

Também observamos que a migração de artistas para outros gêneros e situações trágicas, como assassinatos, também são fatores, mas em menor grau.

O gênero Pop apresentou importante particularidade nessa pesquisa, pois a música mais popular de um ano tende a ser desse gênero. Esse fato foi observado mesmo em períodos onde o Pop não apresenta destaque na quantidade total de músicas.

Enquanto em Araújo et al. [3] buscamos entender como se dá o processo de surgimento de comunidades dentro de um gênero específico, nesse caso o Rap. Para isso, delimitamos a pesquisa a um escopo específico dentro desse cenário, que são as colaborações entre artistas nos álbuns do DJ Khaled, famoso artista americano conhecido por ter sempre vários músicos colaborando em suas músicas (uma delas conta com 12 colaboradores).

Montamos nossa base de dados a partir de uma coleta feita utilizando a API do serviço de *streaming* Napster¹⁰. Feita essa coleta, identificamos os músicos dessa base e montamos representações onde cada um deles era um nó, que eram ligados se haviam participado de uma mesma música. Os pesos dessas arestas foram dados pela quantidade de músicas em que apareciam conjuntamente. Montadas as redes era possível, então, analisá-las. Essa metodologia está apresentada na Figura 2.

Nesse trabalho, montamos nove redes da seguinte forma: a primeira tinha as colaborações do primeiro disco do Khaled, a segunda as referentes ao primeiro e segundo discos e assim sucessivamente. Logo, a nona rede, ilustrada na Figura 3, englobava todas as colaborações do artista no período contemplado na nossa base. Utilizamos a medida de *betweenness centrality* para identificarmos as comunidades destas, assim extraíndo informações de relevância.

Somente foi possível dar significado às comunidades identificadas a partir da rede com as colaborações dos seis primeiros álbuns (essa rede conta com 109 nós e 313 arestas). Ao final, seis comunidades representativas foram identificadas, sendo uma delas a dos novos artistas (em cinza na Figura 3), o que significa que, no caso estudado, eles apresentam maior tendência a colaborar entre si do que com artistas de renome.

Em suma, os principais fatores que identificamos e que estão presentes em nossas publicações são:

⁶<https://twitter.com/>

⁷<https://www.spotify.com/>

⁸<https://www.billboard.com/charts/year-end>

⁹<https://en.wikipedia.org/>

¹⁰<https://us.napster.com/>

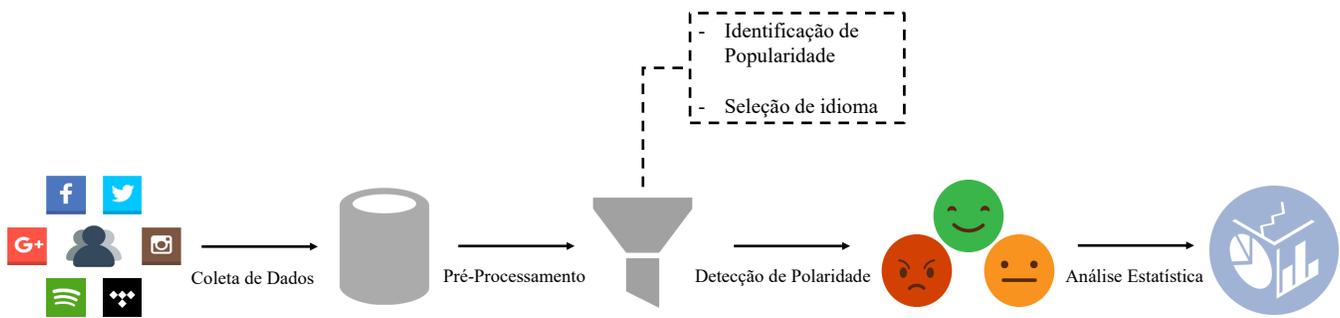


Figura 1: Metodologia utilizada em Araujo et al. [2].

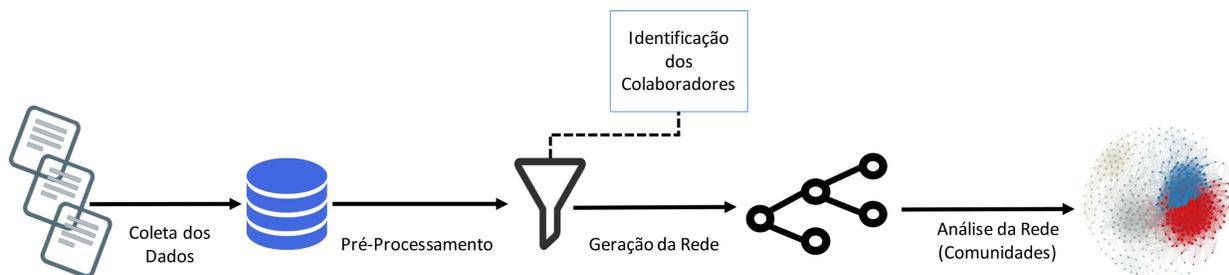


Figura 2: Metodologia utilizada em Araújo et al. [3].

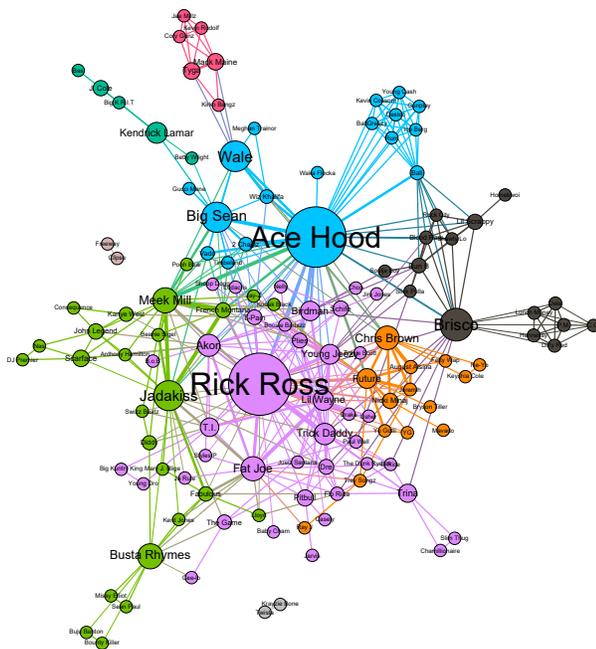


Figura 3: Rede contendo as colaborações entre músicos nos nove primeiros discos do DJ Khaled, extraída de Araújo et al. [3].

- (1) Mensagens com polaridade positiva no Twitter tem impacto sob a popularidade de um álbum no Spotify;
- (2) O surgimento de um ato musical ou um movimento de destaque é o principal fator de influência sob a popularidade de um gênero musical;
- (3) Um artista ou banda podem alavancar sozinhos a popularidade de um gênero por um período de até cinco anos;
- (4) O Pop é o gênero que apresenta maior tendência a ter a música mais popular de um ano;
- (5) No Rap, novos artistas não apresentam tendência a colaborar com artistas consagrados.

4.2 Resultados Esperados

Apesar de já termos obtidos alguns fatores em nossa pesquisa, consideramos que eles ainda não são suficientes para gerarmos um modelo com boa taxa de acerto, pois alguns dos parâmetros estudados tratam apenas de gêneros específicos. Para contornar esse problema devemos buscar generalizar todos os resultados obtidos, mostrando sua validade ou não em um âmbito mais geral.

Além disso, outros fatores ainda necessitam serem estudados. Para que seja possível encontrar entre uma gama de subconjuntos destes, aquele que apresente melhor resultado para utilizarmos em nosso modelo.

Assim, já está planejado um estudo voltado à verbosidade de artistas e sua relação direta com a popularidade destes. Apesar de Marzagão [17] já ter apresentado que quanto maior o vocabulário, menor o sucesso, esse resultado foi apresentado somente para as músicas brasileiras. Além disso, o parâmetro de sucesso difere daquele que planejamos utilizar. Nesse trabalho a colocação nas listas

do *site* Vagalume¹¹ foi usada como indicador, enquanto nós planejamos utilizar o Spotify, que é uma plataforma com abrangência mundial.

Logo, buscaremos uma relação entre o número de palavras distintas utilizadas por um artista em um novo lançamento e o valor em popularidade deste alcançado no Spotify. Podemos gerar um modelo de predição utilizando tal relação como parâmetro único e os analisando estaticamente de forma similar ao que foi feito em Araujo et al. [2].

Também percebemos que álbuns lançados em 2018 por alguns artistas de renome, tais como Beyoncé, Kanye West, Kid Cudi, The Chainsmokers, dentre outros, possuem menos de 10 músicas. Esse fato pode vir a indicar uma nova tendência no mercado musical, buscando adaptar-se às mudanças na forma do consumo de música. Esse fator ainda carece de estudo, onde iremos observar se há relação entre a quantidade de músicas em um álbum e a sua respectiva popularidade. Uma possível correlação entre a duração em minutos e a popularidade também poderia ser obtida como resultado desse estudo.

Com base nesses fatores de influência já obtidos em nossas pesquisas e de estudos publicados por outros pesquisadores, bem como outros que podem ainda vir a ser levantados em nossos estudos, produziremos um modelo de predição, de acordo com a metodologia ilustrada na Figura 4.

O modelo de predição será uma rede neural artificial [25]. O processo de treinamento de uma rede neural artificial é um processo de aprendizado automático no qual obtemos um modelo capaz de representar um conhecimento. Em nosso caso, o conhecimento desejado é a relação entre os fatores previamente estudados e o sucesso ou fracasso de um álbum. Redes neurais artificiais têm sido utilizadas com sucesso para a resolução de diversos problemas que envolvem predição. Por exemplo, para reconhecer atividade humanas [26], realizar diagnósticos clínicos [14] e prever falência de empresas [19].

Um importante aspecto do emprego de estratégias de aprendizado de máquina para tarefas de predição é a definição de um conjunto adequado de atributos de entrada. Assim, iremos estudar a utilidade dos diferentes subconjuntos de fatores (atributos) na geração de modelos preditivos. Consideramos também a possibilidade de empregar modelos de aprendizado profundo [21]. Tais modelos são redes neurais artificiais recorrentes e se destacam por sua capacidade de automaticamente aprender uma representação dos dados originais, que frequentemente produz resultados superiores ao que seriam obtidos pela seleção manual de conjuntos de atributos para representação dos dados. Avaliaremos diversos modelos de predição, os quais serão comparados através de técnicas de validação cruzada [20].

O método a ser proposto fará sua predição para o valor em popularidade no Spotify uma semana após o lançamento do álbum ou *single*.

Em resumo, os próximos passos dessa pesquisa são:

- (1) Generalizar os fatores de influência obtidos que são específicos de um gênero, mostrando sua validade ou não para os outros;

- (2) Identificar outros fatores ao analisar a verbosidade, número de faixas, duração, dentre outros elementos de um álbum;
- (3) Utilizando redes neurais artificiais gerar um modelo que utilize de tais fatores para realizar a predição;
- (4) Validar tal modelo com casos do mundo real.

5 CONCLUSÃO

Apresentamos aqui um projeto de mestrado em andamento que tem como objetivo propôr um algoritmo de predição da popularidade de um álbum no mercado musical. Mais precisamente o valor que esse álbum alcançará em popularidade na plataforma de *streaming* Spotify uma semana após o seu lançamento.

No atual estado de nossa pesquisa estamos analisando possíveis fatores de influência no mercado, visto que esses fatores serão o cerne do modelo que iremos propor.

Em nossos estudos já identificamos que mensagens com popularidade positiva no Twitter influenciam no mercado musical. E, também, que o surgimento de uma banda ou artista de sucesso pode alavancar sozinho a popularidade um gênero por um período de até cinco anos. Mostramos que revoluções musicais são o principal fator de influência sob a popularidade de um gênero. Apresentamos que o gênero Pop possui maior tendência a ter a música mais popular de um ano e que artistas novos não tendem a colaborar com os já consagrados no Rap [1] [2] [3].

Os próximos passos são elencar ainda mais fatores e utilizando-se de redes neurais artificiais, buscar um subconjunto de fatores e pesos que entregue o melhor modelo de predição. Esse modelo gerado será testado e validado com casos do mundo real, onde buscaremos atingir a maior taxa de acerto possível. Assim, estabelecendo um valor base para estudos futuros nessa área.

AGRADECIMENTOS

Os autores agradecem ao CNPq pelo financiamento dessa pesquisa.

REFERÊNCIAS

- [1] Carlos V.S. Araujo and Eduardo F. Nakamura. 2018. Identification of Most Popular Musical Genres and their Influence Factors. In *Proceedings of the 24th Brazilian Symposium on Multimedia and the Web (WebMedia '18)*. ACM, New York, NY, USA, 4. <https://doi.org/10.1145/3243082.3264665>
- [2] Carlos V.S. Araujo, Rayol M. Neto, Fabiola G. Nakamura, and Eduardo F. Nakamura. 2017. Predicting Music Success Based on Users' Comments on Online Social Networks. In *Proceedings of the 23rd Brazilian Symposium on Multimedia and the Web (WebMedia '17)*. ACM, New York, NY, USA, 149–156. <https://doi.org/10.1145/3126858.3126885>
- [3] Carlos V.S. Araujo, Rayol M. Neto, Fabiola G. Nakamura, and Eduardo F. Nakamura. 2017. Using Complex Networks to Assess Collaboration in Rap Music: A Study Case of DJ Khaled. In *Proceedings of the 23rd Brazilian Symposium on Multimedia and the Web (WebMedia '17)*. ACM, New York, NY, USA, 425–428. <https://doi.org/10.1145/3126858.3131605>
- [4] Pascal Budner and Joern Grahl. 2016. Collaboration Networks in the Music Industry. *arXiv preprint arXiv:1611.00377* (2016).
- [5] Conny Chen and Ramnath Chellappa. 2009. Music sales and online user activity at social networking sites. In *fifth symposium of the Heinz College's iLab Center for Digital Media Research, Carnegie Mellon University on Statistical Challenges in Electronic Commerce Research, Pittsburgh*. 30–31.
- [6] Hailiang Chen, Prabuddha De, and Yu Jeffrey Hu. 2015. IT-enabled broadcasting in social media: An empirical study of artists' activities and music sales. *Information Systems Research* 26, 3 (2015), 513–531.
- [7] Sanjeev Dewan and Jui Ramaprasad. 2012. Research Note—Music Blogging, Online Sampling, and the Long Tail. *Information Systems Research* 23, 3-part-2 (2012), 1056–1067. <https://doi.org/10.1287/isre.1110.0405>
- [8] Sanjeev Dewan and Jui Ramaprasad. 2014. Social media, traditional media, and music sales. *Mis Quarterly* 38, 1 (2014), 101–121.

¹¹<https://www.vagalume.com.br/>

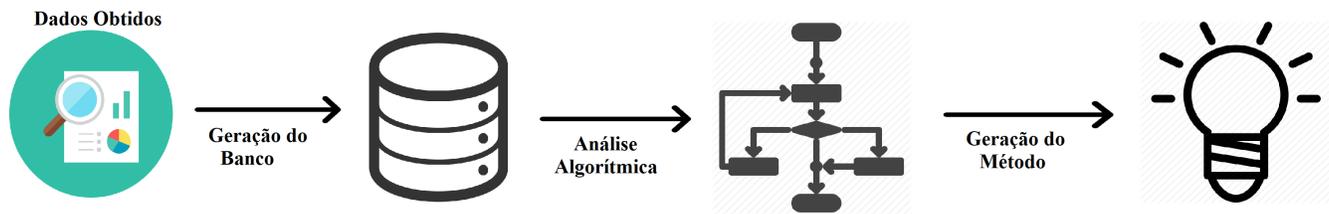


Figura 4: Metodologia proposta para a geração do método de predição.

- [9] Vasant Dhar and Elaine A. Chang. 2009. Does Chatter Matter? The Impact of User-Generated Content on Music Sales. *Journal of Interactive Marketing* 23, 4 (2009), 300 – 307. <https://doi.org/10.1016/j.intmar.2009.07.004>
- [10] Robert Dimery. 2007. 1001 discos para ouvir antes de morrer. *Tradução Carlos Irineu da Costa e Eliane Azevedo. Rio de Janeiro: Sextante* (2007).
- [11] Gabriele Giaquinto, Cora Bledsoe, and Brian McGuirk. 2007. Influence and similarity between contemporary jazz artists, plus six degrees of kind of blue. *Ph. D. Dissertation, Citeseer* (2007).
- [12] Pablo M Gleiser and Leon Danon. 2003. Community structure in jazz. *Advances in complex systems* 6, 04 (2003), 565–573.
- [13] David Grandadam. 2008. Evolving networks and the finest in jazz. In *Creative Industries and Intellectual Property Conference*, Vol. 22.
- [14] Zachary Chase Lipton, David C. Kale, Charles Elkan, and Randall C. Wetzell. 2015. Learning to Diagnose with LSTM Recurrent Neural Networks. *CoRR abs/1511.03677* (2015). arXiv:1511.03677 <http://arxiv.org/abs/1511.03677>
- [15] Paul D. Lopes. 1992. Innovation and Diversity in the Popular Music Industry, 1969 to 1990. *American Sociological Review* 57, 1 (1992), 56–71. <http://www.jstor.org/stable/2096144>
- [16] Rayol M Neto, Bruno Ábia, Thais Almeida, Fabiola Guerra Nakamura, and Eduardo Nakamura. 2017. Uma Abordagem para Identificação de Entidades Influentes em Eventos Comentados nas Redes Sociais Online. In *Proceedings of the 14th Simpósio Brasileiro de Sistemas Colaborativos*.
- [17] Thiago Marzagão. 2015. text-analyzing Brazilian music. (2015). <http://thiogomarzagao.com/2015/12/24/text-analyzing-brazilian-music/>
- [18] Matthias Mauch, Robert M. MacCallum, Mark Levy, and Armand M. LeROI. 2015. The evolution of popular music: USA 1960–2010. *Royal Society Open Science* 2, 5 (2015). <https://doi.org/10.1098/rsos.150081> arXiv:<http://rsos.royalsocietypublishing.org/content/2/5/150081.full.pdf>
- [19] M. D. Odom and R. Sharda. 1990. A neural network model for bankruptcy prediction. In *1990 IJCNN International Joint Conference on Neural Networks*, 163–168 vol.2. <https://doi.org/10.1109/IJCNN.1990.137710>
- [20] Stuart J Russell and Peter Norvig. 2010. *Artificial Intelligence: a Modern Approach*. Pearson Education Inc.
- [21] Jürgen Schmidhuber. 2015. Deep learning in neural networks: An overview. *Neural Networks* 61 (2015), 85 – 117. <https://doi.org/10.1016/j.neunet.2014.09.003>
- [22] Reginald D Smith. 2006. The network of collaboration among rappers and its community structure. *Journal of Statistical Mechanics: Theory and Experiment* 2006, 02 (2006), P02006.
- [23] Bruno A Souza, Thais G Almeida, Alice AF Menezes, Fabiola G Nakamura, Carlos Figueiredo, and Eduardo F Nakamura. 2016. For or against?: Polarity analysis in tweets about impeachment process of brazil president. In *Proceedings of the 22nd Brazilian Symposium on Multimedia and the Web*. ACM, 335–338.
- [24] Brian Uzzi. 2008. Collaboration Networks and Innovation: The Problem of Clustering. (2008).
- [25] Marcel van Gerven and Sander Bohte. 2017. Editorial: Artificial Neural Networks as Models of Neural Information Processing. *Frontiers in Computational Neuroscience* 11 (2017), 114. <https://doi.org/10.3389/fncom.2017.00114>
- [26] Jianbo Yang, Minh Nhut Nguyen, Phyo Phyo San, Xiaoli Li, and Shonali Krishnaswamy. 2015. Deep Convolutional Neural Networks on Multichannel Time Series for Human Activity Recognition.. In *Ijcai*, Vol. 15. 3995–4001.

WebMedia 2018
XXIV Simpósio Brasileiro de Sistemas Multimídia e Web

ANAIS

**XV Workshop de Trabalhos de
Iniciação Científica**

Coordenadores

Tiago Maritan (UFPB)
Antônio Maurício Pitangueira (IFBA)
Antônio Carlos Souza (IFBA)

XV Workshop de Trabalhos de Iniciação Científica (WTIC)

Sumário

Prefácio	46
Organização	47
Sessão Técnica 1: Sistemas Web, Redes Sociais e IoT	49
WALO: Ferramenta para Gerenciamento de Mídias Sociais	49
Marília G. Santos (IFBA), Antonio C.S. Souza (IFBA), Márcio C.F. Macedo (UFBA)	
Recuperação de Imagens na Web com Fusão Adaptativa de Credibilidade Baseada em Algoritmos Genéticos	53
Wanderson Bezerra da Silva (UEFS), Rodrigo Tripodi Calumby (UEFS)	
Using images to extend smart object discovery in an Internet of Things scenario	57
Romulo Silva (UFC), Windson Viana (UFC), Paulo Filipe Dantas (UFC)	
EasyContext: Facilitando o desenvolvimento de aplicações com a Awareness API	61
Davi Batista Tabosa (UFC), Windson Viana (UFC)	
Sessão Técnica 2: Sistemas, Aplicações Multimídia e Interação	65
Towards Data Transmission Through Inaudible Sound in Ginga-NCL	65
João Victor G. de S. Nunes (PUC-Rio), Álan L.V. Guedes (PUC-Rio), Guilherme F. Lima (PUC-Rio), Sérgio Colcher (PUC-Rio)	
MathQuiz: A Game App for M-Learning	69
José Marcelo Silva Santiago (UECE), Monck Charles Nunes de Albuquerque (UECE), Fernanda Kécia de Almeida (UECE), Francisco Ranulfo Freitas Martins (UECE), Yuri Soares de Oliveira (IFCE)	
Assessment of the Treatment of Grammatical Aspects of Machine Translators to Libras	73
Luana Silva Reis (UFPB), Tiago Maritan U. de Araújo (UFPB), Yuska Paola Costa Aguiar (UFPB), Manuella Aschoff C.B. Lima (UFPB), Angelina S. da Silva Sales (UFPB)	
Reformulação do Projeto de Usabilidade da Ferramenta Lagarto	77
Samir Chaves (UFC), Fernando Trinta (UFC)	
Levantamento de Tecnologias para ONGs de Proteção Animal para apoio ao resgate de animais domésticos acoplados ao ciclo de vida de um sistema Web	81
Tatiana Tozzi (IFC), Daniel Fernando Anderle (IFC), Rodrigo Ramos Nogueira (IFC)	

XV Workshop de Trabalhos de Iniciação Científica

Prefácio

O XV Workshop de Trabalhos de Iniciação Científica (WTIC) é um dos eventos integrantes do Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia) e constitui um importante fórum dedicado à apresentação e discussão de pesquisas em desenvolvimento por estudantes de graduação nas áreas de Multimídia, Hipermídia e Web. O evento visa dar visibilidade e integrar os potenciais pesquisadores junto às comunidades acadêmica e industrial.

Neste ano, o WTIC recebeu um total de 16 submissões de artigos técnicos de diversas instituições do país e em todas as áreas de interesse do WebMedia. Os artigos foram avaliados de forma criteriosa por um Comitê de Programa (CP) de alto nível, o que levou à seleção de 10 trabalhos de iniciação científica para apresentação e publicação nos anais do evento (resultando assim em uma taxa de aceitação de 62,5%). A alta taxa de aceitação reflete o principal objetivo do evento de atrair alunos de graduação interessados na área do evento.

Aproveitamos este espaço também para agradecer a todos os autores que submeteram artigos para o Workshop, contribuindo assim para o seu enriquecimento, aos membros do comitê de programa por sua dedicação, pelas valiosas revisões dos artigos e pelo comprometimento com os prazos estabelecidos. Ao Professor Tiago Maritan Ugulino de Araújo (UFPB) e aos Professores Antonio Carlos Souza e Antonio Maurício Pintagueira (IFBA) pelo aceite ao convite para coordenar o WTIC e por todo intenso trabalho que vêm desenvolvendo na organização do evento, e à CE-WebMedia da SBC pela confiança depositada no nosso trabalho.

Salvador, outubro de 2018.

Tiago Maritan (UFPB)

Antônio Maurício Pitagueira (IFBA)

Antônio Carlos Souza (IFBA)

Coordenadores do XV Workshop de Trabalhos de Iniciação Científica

Coordenadores do XV WTIC

Tiago Maritan (UFPB)
Antônio Maurício Pitangueira (IFBA)
Antônio Carlos Souza (IFBA)

Comitê de Programa do XV WTIC

Alan Guedes (PUC-Rio)
Carlos Eduardo Batista (UFPB)
Carlos Soares Neto (UFMA)
César Teixeira (UFSCar)
Eduardo Barrére (UFJF)
José Valentim dos Santos Filho (UFRB)
Lynn Alves (UFBA)
Marcelo Moreno (UFJF)
Márcio Macêdo (UFBA)
Rostand Costa (UFPB)
Tatiana Tavares (UFPel)
Thaís Gaudêncio do Rêgo (UFPB)

WALO: Ferramenta para Gerenciamento de Mídias Sociais

Marília G. Santos
Instituto Federal da Bahia
R. Emídio dos Santos, Barbalho
Salvador, Bahia 40.301-015
mariliasantos@ifba.edu.br

Antonio C. S. Souza
Instituto Federal da Bahia
R. Emídio dos Santos, Barbalho
Salvador, Bahia 40.301-015
antoniocarlos@ifba.edu.br

Márcio C. F. Macedo
Universidade Federal da Bahia
Av. Adhemar de Barros, Ondina
Salvador, Bahia 40.170-110
marcio.cerqueira@ufba.br

ABSTRACT

Social media is a very important tool for businesses, especially those that depend on marketing strategies for greater brand recognition. In addition, social media helps enrich the customer experience and reduce marketing expenses. Social media management allows companies to get information that reflects the overall consumer sentiment of their brand or their products, quickly and securely. However, such management is not an easy task, since the large amount of existing social media requires a lot of effort from professionals to deal with all the information that influences these social media. Therefore, a reasonable need is identified in creating a specific tool to manage social media at once. In this article, we propose WALO, a new open source tool that allows professionals to use different social media simultaneously, providing a new way for these professionals to interact with their clients through a chatbot.

KEYWORDS

Social networks, social media management, open-source tool

1 INTRODUÇÃO

As mídias sociais possuem muito conteúdo, em virtude da troca de informações que é realizada a todo momento entre os usuários. As empresas estão gradativamente utilizando as mídias sociais para divulgar sua marca, porque assim é possível ter maior visibilidade no mercado com baixo custo, ou mesmo custo zero de investimento, já que as mídias sociais são ambientes propícios para a interação dos usuários com a marca [9].

Atualmente, há um custo elevado para construir uma marca com alta reputação na *Internet*, uma vez que é necessário planejamento estratégico de comunicação e *marketing* de modo que a marca se torne conhecida e tenha credibilidade. Eventualmente, durante esse processo, podem surgir solicitações e comentários desfavoráveis nas mídias sociais a respeito da empresa, devido às exigências dos consumidores, que sempre querem maiores detalhes das informações disponíveis a respeito da marca com menor tempo de resposta possível [5].

As empresas que planejam entrar no mundo virtual devem realizar um estudo, fazendo o levantamento das reais necessidades dos potenciais consumidores, a fim de transmitir segurança quanto à empresa e proporcionar uma melhor impressão sobre marca. Assim, tendo em vista que as empresas precisam acompanhar a sua

reputação nas mídias sociais, surge a necessidade do desenvolvimento de uma ferramenta com interface intuitiva e que viabilize o monitoramento da reputação da empresa de acordo com as manifestações (postagens, notificações, possíveis dúvidas) dos usuários nas mídias sociais.

De acordo com o exposto, propõe-se neste artigo uma ferramenta *open-source* para gerenciamento de mídias sociais que almeja possibilitar a comunicação de uma empresa com os seus potenciais clientes através de um *chatbot*, que terá como funções principais a análise dos comentários nas mídias sociais e a interação com os seus usuários, podendo então sanar dúvidas em comum, otimizando e diminuindo a quantidade de atendimentos realizados pelos colaboradores da empresa. Dessa forma, a ferramenta WALO permitirá o acompanhamento e a análise dos questionamentos realizados pelos clientes.

O restante deste artigo está estruturado da seguinte forma:

- A Seção 2 contém uma breve descrição dos trabalhos relacionados ao tema deste artigo. Nessa Seção é feita uma comparação entre o WALO e as ferramentas existentes tanto na literatura quanto na academia para o gerenciamento das mídias sociais;
- A Seção 3 apresenta a descrição da ferramenta *open-source* para gerenciamento de mídias sociais que é proposta neste artigo;
- Por fim, na Seção 4, são apresentadas as considerações finais do trabalho, bem como uma perspectiva de trabalhos futuros para o desenvolvimento e avaliação da ferramenta;

2 TRABALHOS RELACIONADOS

Nesta seção, serão apresentados alguns trabalhos relacionados ao tema proposto para o artigo.

Almeida *et al.* [2] abordam uma análise do impacto de diversos tipos de conteúdos acerca do *marketing viral*. Estas são técnicas de *marketing* que têm em vista explorar as mídias sociais com o propósito de obter uma maior repercussão da marca. Os autores utilizaram o *Facebook* para analisar oito perfis de marcas brasileiras de cerveja e seu impacto na rede social. Este artigo é relevante para o campo de análise de mídias sociais, uma vez que demonstra a influência das postagens em mídias sociais e no *marketing* das empresas.

Kim *et al.* [7] expressam uma visão geral do impacto da influência social e da tomada de decisões sobre o comércio eletrônico para conduzir os pesquisadores e empresas que têm interesse em questões relacionadas às mídias sociais. É possível capturar comportamentos dos clientes, da mesma maneira que a influência social pode ser usada pelo comércio eletrônico, e sites podem ser utilizados

In: XV Workshop de Trabalhos de Iniciação Científica (WTIC 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pôsteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.

© 2018 SBC – Sociedade Brasileira de Computação.
ISBN 978-85-7669-435-9.

para auxiliar no processo de tomada de decisão dos potenciais consumidores. Os autores discutem que é necessário acompanhar a repercussão da marca nas mídias sociais e a importância de cada produto ou serviço nos comentários, tornando possível o acompanhamento imediato do que cada cliente pode estar comentando sobre a marca, o que pode influenciar no alcance de potenciais clientes.

Victória Aires, Fabiola Nakamura e Eduardo Nakamura [1] abordam o aproveitamento das redes sociais, tendo em vista que representam relações tanto de amizade quanto de trabalho. Os autores estudam se as redes sociais tem influência no sucesso de um filme. Para esse estudo, eles realizaram a análise das redes sociais de cinco filmes, a fim de apresentar resultados que comprovassem que o sucesso de bilheteria de um filme está associado ao sucesso que ele faz na rede social. Apesar dos resultados terem sido positivos, o estudo realizado consistiu apenas em uma análise introdutória acerca do tema, tendo em vista que há a necessidade da coleta de uma maior base de dados, possibilitando uma análise mais detalhada dos filmes em suas redes sociais.

Na indústria, já existem algumas ferramentas que auxiliam os usuários no gerenciamento de suas mídias sociais.

HootSuite [6] é uma ferramenta que auxilia a realização do gerenciamento das mídias sociais em uma única interface, viabilizando a criação de várias páginas de conteúdo personalizável. Em cada aba referente a uma mídia social, é possível adicionar diferentes colunas com o que é postado e recepcionado. Através do *HootSuite*, é possível programar a publicação de postagens, definindo a data, horário e o perfil em que se deseja realizar a publicação.

A ferramenta é bastante utilizada para o gerenciamento de vários perfis em uma mesma rede, possibilitando assim a realização de uma análise geral dos perfis, sem que seja necessário efetuar *login/logout* para cada conta. Sua interface é simples e proporciona agilidade e pleno controle do que acontece nas mídias sociais, permitindo que seja avaliado como, por quem e quando a informação for recebida.

Esta ferramenta é a mais próxima da proposta deste trabalho, porém possui três desvantagens: (i) a *interface* é confusa para quem necessita gerenciar vários perfis da mesma mídia social; (ii) a versão gratuita só permite o gerenciamento de até três perfis; (iii) é necessária a realização de um pagamento para se obter um relatório de análise mais detalhado. Enquanto isso, a WALO se propõe a ser uma ferramenta *open-source*, que não oferece nenhum custo ao usuário, assim como possui uma *interface* intuitiva, de fácil compreensão e visualização das publicações no painel de monitoramento.

Postcron [8] é uma das ferramentas mais conhecidas no mercado quando o assunto é agendamento de publicações no *Facebook*. É um aplicativo que permite agendamento de forma rápida, em várias mídias sociais. Outra grande vantagem é a sua extensão integrada ao *Chrome*, que facilita o agendamento de publicações, porém são limitadas às possíveis mídias sociais que podem ser integradas, assim como há um custo alto a ser investido para aquisição, e os valores variam de acordo com a quantidade de contas, postagens e quantidade de analistas de mídias sociais que gerirão a ferramenta.

Buffer [4] é uma ferramenta para gerenciamento de mídias sociais voltada para os analistas da área. Essa ferramenta possui diversas funcionalidades para auxiliar no gerenciamento dos perfis, além disso, a ferramenta oferece uma variedade de aplicativos *mobile*

e extensões de navegador, possibilitando adicionar conteúdo ao mesmo tempo em que os analistas de mídias sociais vão descobrindo o perfil. O *Buffer* é uma ferramenta intuitiva, o que torna fácil disseminar conteúdos nas mídias sociais, porém nas versões em que há mais funcionalidades existem limitações, uma vez que a ferramenta não é *open source*. A ferramenta não permite interação direta com os utilizadores, e sua versão gratuita tem recursos bastante limitados.

Tendo em vista os trabalhos e ferramentas listados, a WALO, além de realizar o gerenciamento das mídias sociais, tem como diferencial a varredura nos comentários realizados pelos potenciais clientes, de modo que as dúvidas possam ser sanadas pelo *botWalo*, que será inicializado a partir do reconhecimento de alguma palavra informada nos comentários, que já estará cadastrada na base de conhecimento do *bot*, podendo assim diminuir a quantidade de acionamentos a um humano.

3 WALO

O objetivo da WALO é realizar, através da ferramenta *open-source*, o gerenciamento de mídias sociais, de forma que a própria comunidade possa agregar novas funcionalidades no projeto de forma extensível e transparente. A plataforma visa facilitar o gerenciamento e monitoramento das marcas de diferentes empresas nas mídias sociais, sanando o maior número possível de questionamentos publicados por potenciais clientes, de forma que a empresa possa prestar melhor suporte e retorno mais rápido para o cliente. A partir dessa interação rápida e eficaz, em que a empresa esclarece a dúvida do cliente, a marca da empresa é fortalecida no mercado, pois quanto menor o tempo de resposta de cada solicitação, maior credibilidade as empresas ganham frente às mídias sociais.

A finalidade do uso do *chatbot* na ferramenta é auxiliar o usuário no esclarecimento de suas dúvidas. Para tanto, o *chatbot* realiza seu processamento baseado em palavras já cadastradas numa base de dados. O *chatbot* tem como propósito diminuir a necessidade de acionamento de um humano para responder questionamentos que se repetem nas mídias sociais. Porém, o cliente ainda poderá solicitar contato com um colaborador da empresa a qualquer momento, durante a interação com o *botWalo*. Isso porque, na WALO, o *chatbot* não se passará por um humano, apenas tentará esclarecer a dúvida do cliente de forma rápida e objetiva.

Além do que já foi exposto, a ferramenta possibilita que o usuário realize uma nova postagem na WALO e visualize essa postagem na mídia social escolhida, bem como a busca das postagens.

Para o desenvolvimento da ferramenta, foram escolhidas duas redes sociais: *Facebook* e *Twitter*, já que elas fazem parte das redes sociais mais acessadas no mundo [3].

Para auxiliar na implantação do *chatbot* na WALO foi utilizado o *Dialogflow*¹, que é um pacote de desenvolvimento da *Google*, criado para desenvolver *interfaces* de tecnologias de interação humano-computador baseadas em recursos de processamento de linguagem natural, que serão adicionadas aos aplicativos, serviços e dispositivos. Fazendo uso da tecnologia de aprendizagem de máquina, o *Dialogflow* reconhece a intenção e o contexto da fala de um usuário, de modo que na comunicação o *bot* possa oferecer respostas eficientes e precisas.

¹<https://dialogflow.com/>

As áreas de conhecimento do *BotWalo* foram definidas em duas possibilidades: a do conhecimento específico e a do conhecimento adicional. A área de conhecimento específico possui informações já cadastradas, vinculadas aos produtos e serviços expostos nas redes sociais. A base de conhecimento foi extraída dos comentários de clientes nas redes sociais antes do desenvolvimento da ferramenta. O *BotWalo* não foi desenvolvido com o propósito de se passar por um humano, por isso a área de conhecimento adicional é captada pelo usuário administrador da ferramenta, através dos dados estatísticos extraídos da ferramenta, notificando a quantidade de chamados atendidos pelo humano e a quantidade de chamados atendidos através do *BotWalo*. Caso a quantidade de atendimentos realizados pelo *BotWalo* seja maior que aquela realizada pelo humano, a área de conhecimento do *BotWalo* está atendendo às expectativas, caso contrário será necessário adicionar informações ao *bot*, de modo que ele esteja o mais completo possível e consiga sanar o maior número de questionamentos realizados pelos clientes.

3.1 Levantamento de Requisitos

Nesta seção são apresentadas todas as condições que foram definidas como requisitos funcionais e não funcionais relevantes para a construção da ferramenta proposta.

3.1.1 Requisitos Funcionais. Através dos requisitos funcionais são descritas as funcionalidades e serviços desenvolvidos no sistema. Na tabela 1 são listados os requisitos funcionais.

3.1.2 Requisitos não funcionais. O conjunto de requisitos não funcionais do sistema estão listados logo abaixo.

Usabilidade. A ferramenta proposta deve ser simples e de fácil utilização;

Manutenibilidade. A ferramenta proposta deve possibilitar reparos de forma facilitada, bem como a inclusão de novas funcionalidades, visando assim o aperfeiçoamento do *software*.

Desempenho. A plataforma desenvolvida deve buscar identificar, nos possíveis comentários dos clientes, dúvidas que possam ser sanadas através do *chatbot*, que verifica periodicamente as novas mensagens, a fim de proporcionar maior rapidez no tempo de resposta.

Portabilidade. Por se tratar de um sistema *web*, a ferramenta desenvolvida deve ser portátil, podendo ser utilizada em navegadores e sistemas operacionais distintos.

Segurança. A ferramenta deve oferecer segurança por meio de autenticação ao sistema;

3.2 Arquitetura do Sistema

A arquitetura da WALO é formada por três camadas, explicitando o padrão arquitetural MVC (*Model-View-Controller* ou Modelo-Controle-Visão), o qual viabiliza a independência dos dados trabalhados e das ações do moderador que necessita desses dados.

A camada *Model* fica encarregada da lógica e inteligência do sistema, da validação de negócio e processamento. A camada *View* se encarrega da parte da aplicação que interage com o usuário moderador, nesta camada é dada visibilidade para os componentes na apresentação dos dados. Já a camada *Controller* se encarrega

Table 1: Requisitos funcionais estabelecidos para o WALO.

Nº	Requisitos
RF1	O Usuário deve ser capaz de monitorar pelo menos uma mídia social através ferramenta;
RF2	A ferramenta tem apenas um perfil de usuário moderador, o qual deve moderar e liderar as possíveis discussões realizando todo o acompanhamento das mídias sociais;
RF3	O usuário autenticado deve visualizar e responder aos comentários dos potenciais clientes;
RF4	Cada mídia social monitorada deve exibir as postagens com suas respectivas curtidas e comentários;
RF5	O sistema deve verificar as postagens dos clientes, de modo a identificar possíveis dúvidas e questionamentos;
RF6	O sistema deve exibir a postagem inicialmente em formato sintetizado;
RF7	O sistema deve possibilitar a interação com as mídias sociais através do <i>chatbot</i> quando houver algum questionamento acerca dos seus projetos;
RF8	O sistema deve permitir que o moderador responda aos questionamentos dos clientes nas redes sociais, caso não obtenha resposta através do <i>chatbot</i> ;
RF9	O sistema deve exibir gráficos com a quantidade de solicitações respondidas pelo <i>chatbot</i> e avaliação do <i>chatbot</i> pelo humano;
RF10	O moderador poderá consultar a interação do <i>bot</i> com as mídias sociais a cerca dos seus produtos;

de receber as requisições do usuário moderador, a qual interage com a *view* que monta os objetos correspondentes e passa-os à camada *Model* para executar a solicitação. A WALO possibilita a incorporação de novas redes sociais, através da chamada do método *require* na classe *Monitoramento*. A Figura 1 apresenta a arquitetura do *botWalo*. Através dele os usuários podem ter respostas aos seus questionamentos de forma mais rápida e objetiva. O sistema principal é expresso através da compreensão da linguagem, acesso e recuperação das informações, e do controlador de conversões. O sistema também é conectado através de uma fonte de dados, que realiza armazenamento das respostas para as perguntas mais frequentes, popularmente conhecidas como FAQs (*Frequently Asked Questions*).

Language Understanding. O desafio associado ao uso do *botWalo* está relacionado à área de PLN (Processamento de Linguagem Natural), uma vez que uma limitação da identificação das palavras mais utilizadas nas redes sociais impede que *botWalo* seja acionado de forma apropriada para interação com o usuário. Para tanto, FAQs serão armazenadas numa base de dados, para que o *bot* consiga identificar as perguntas e responder seus cliente. Para os questionamentos que não podem ser respondidos através do *botWalo*, também há armazenamento no banco de dados do registro de conversas, de modo que sejam analisados para inserção de novas FAQs. Isso permite que haja uma diminuição das solicitações dos clientes para realizar contato com um humano.

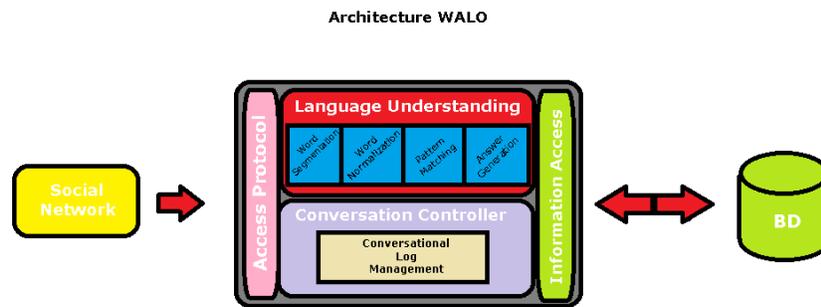


Figure 1: Arquitetura Model Viewer Controller

Conversation Controller. O controlador de conversão fornece a qualidade de serviço adequada para as sessões, bem como permite que se entenda o que cada cliente questiona e que se mantenha os arquivos de *log* da conversação. Esta função consiste em gerenciar os logs de conversas onde há todos os registros de conversas entre os usuários e o sistema. Através desses registros é possível analisar questões específicas caso o sistema não esteja fornecendo algum tipo de resposta. Além disso, este módulo fornece algumas estatísticas: (i) número de usuários que deram *like* numa postagem, (ii) quantidade de acionamentos do *botWalo*, (iii) solicitações de direcionamento para falar com um humano, (iv) quantidade perguntas ainda desconhecidas pelo *botWalo*.

4 CONCLUSÃO E TRABALHOS FUTUROS

Neste artigo, foi apresentada a proposta da WALO: uma ferramenta para gerenciamento de mídias sociais, que tem como finalidade auxiliar os analistas da área no monitoramento das redes sociais, bem como, na busca de possíveis questionamentos publicados. A fim de esclarecer estas dúvidas, almeja-se o uso de um *chatbot* na WALO. Com o intuito de ajudar os profissionais que utilizam as mídias sociais para divulgação da marca, a plataforma disporá de um processo centralizado e automatizado na busca de possíveis questionamentos dos potenciais clientes. Visando otimizar a centralização dos questionamentos dos clientes em uma única plataforma, já que as discussões ocorrem em mídias distintas, este artigo mostra uma possível solução para agregar várias mídias sociais em uma plataforma.

Como trabalhos futuros, almeja-se o desenvolvimento e lançamento da ferramenta WALO para o público em geral, adicionando o suporte ao monitoramento de várias mídias sociais, o suporte à postagem de conteúdos multimídias nas mídias sociais de forma programada, bem como a integração dos *chatbots* à plataforma, para auxílio no controle da reputação da marca da empresa na *Internet*.

REFERÊNCIAS

- [1] Victória P. Aires, Fabíola G. Nakamura, and Eduardo F. Nakamura. 2017. A Social Network Analysis of Successful Movies with Complex Networks. In *Proceedings*

of the 23rd Brazillian Symposium on Multimedia and the Web: Workshops and Posters. 99–102.

- [2] Marcos Inácio Severo de Almeida, Milena Costa, Ricardo Limongi Franca Coelho, and Paulo Roberto Scalco. 2016. "Engaja-me e atraia-me, então eu compartilharei": uma análise do impacto da categoria da postagem no marketing viral em uma rede social. *Revista Brasileira de Gestão de Negócios* 18 (12 2016), 545 – 569.
- [3] Fernanda Beling. 2016. As 10 maiores redes sociais. (2016). <https://www.oficinadanet.com.br/post/16064-quais-sao-as-dez-maiiores-redes-sociais> [Online; accessed 18-julho-2018].
- [4] Buffer. 2018. Buffer. <https://buffer.com/>. (2018). Accessed: 2018-07-30.
- [5] Yany Grégoire, Audrey Salle, and Thomas M Tripp. 2015. Managing social media crises with your customers: The good, the bad, and the ugly. *Business Horizons* 58, 2 (2015), 173–182.
- [6] HootSuite. 2018. HootSuite. <https://hootsuite.com>. (2018). Accessed: 2018-07-30.
- [7] Young Ae Kim and Jaideep Srivastava. 2007. Impact of Social Influence in e-Commerce Decision Making. In *Proceedings of the Ninth International Conference on Electronic Commerce*. ACM, New York, NY, USA, 293–302.
- [8] PostCron. 2018. PostCron. <https://postcron.com>. (2018). Accessed: 2018-07-30.
- [9] Cong Qi and Patrick YK Chau. 2016. An Empirical Study of the effect of Enterprise Social Media Usage on Organizational Learning. In *Proceedings of PACIS*. 330.

Recuperação de Imagens na Web com Fusão Adaptativa de Credibilidade Baseada em Algoritmos Genéticos

Wanderson Bezerra da Silva
Universidade Estadual de Feira de Santana
Avenida Transnordestina, s/n - Novo Horizonte
Feira de Santana, Bahia 44036-900
wbsilva@ecompu.uefs.br

Rodrigo Tripodi Calumby
Universidade Estadual de Feira de Santana
Avenida Transnordestina, s/n - Novo Horizonte
Feira de Santana, Bahia 44036-900
rtcalumby@uefs.br

ABSTRACT

Credibility information gives an indication of which users are most likely to share relevant images on social network feeds and consequently may help estimating the relevance of an image for retrieval purposes. Considering multiple credibility evidences has been shown as an effective method for image ranking. In order to select and combine multiple credibility descriptors, this work proposes a genetic algorithm-based automatic context-adaptive weight adjustment model. The experimental results show promising effectiveness when compared to the baseline.

KEYWORDS

Machine Learning, Image Retrieval, Credibility Descriptors

1 INTRODUÇÃO

Dadas as necessidades dos usuários e a grande quantidade de informação disponível na web, é imprescindível o uso de técnicas eficientes e eficazes para exploração de grandes coleções de dados. Tratando-se de busca de imagens, por exemplo, a abordagem mais comum baseia-se na utilização de informações textuais (metadados, palavras-chaves, páginas web, etc.) [1]. Outro paradigma utiliza a descrição das propriedades visuais (cor, forma, textura, estrutura, etc.) das imagens para indexá-las [4]. Quando consideramos o compartilhamento de informação em mídias sociais, outro fator importante é a reputação das fontes de dados, que podem representar a qualidade e confiabilidade do conteúdo [8].

Neste contexto, informações de credibilidade considerando as tags que um usuário associa ao conteúdo das imagens que compartilha, buscam fornecer uma estimativa automática da qualidade global do conteúdo presente nas relações tag-imagem para as contribuições deste usuário [8]. A credibilidade de um usuário pode ser computada levando-se em consideração fatores como, e.g., a descrição textual da imagem, a quantidade de visualizações, frequência de coocorrência entre as tags de um usuário e entre diferentes usuários, proporção de imagens que apresentam faces humanas, similaridade entre as tags de localização das imagens que representam o mesmo local, entre outros. Estas informações são particularmente valiosas para explorar o contexto social dos dados. Elas fornecem uma indicação sobre quais usuários são mais propensos a compartilhar imagens com anotações relevantes/confiáveis. Neste processo, a

credibilidade do usuário é um fator de grande importância para o refinamento dos resultados obtidos por um sistema de busca. Contudo, um único critério de credibilidade pode não ser capaz de capturar adequadamente a reputação de um usuário e permitir que suas imagens tenham seu grau de relevância corretamente estimado para um determinado contexto de busca. De modo a atenuar este problema, descritores de credibilidade podem ser combinados para refinar os resultados da busca, e.g., eliminando-se itens fornecidos por usuários de baixa credibilidade. Esta não é uma tarefa fácil, dado que é preciso definir quais as melhores combinações entre os descritores de modo a prover o melhor resultado para um determinado cenário de busca.

O uso dos descritores de credibilidade tem atraído grande atenção de pesquisadores nos últimos anos. Neste sentido, podemos citar a MediaEval Retrieving Diverse Social Images Task [7], desafio que teve por cenário a simulação do uso de mecanismos de buscas em redes sociais por um turista que está interessado em obter uma visão geral sobre uma determinada localidade, por exemplo, pontos turísticos e regiões geográficas, antes de escolher seu destino. Nesta atividade, descritores de credibilidade podem ser utilizados, e.g., para ranquear, filtrar e/ou diversificar os resultados em sistemas de recuperação de imagens.

Considerando as limitações individuais dos descritores de credibilidade e a necessidade de modelos eficazes para sua combinação, este trabalho propõe uma técnica de fusão linear para descritores de credibilidade, cujos pesos são automaticamente ajustados por meio da meta-heurística de Algoritmos Genéticos (AG). Assim, estes algoritmos são utilizados para a descoberta da importância/contribuição de cada descritor, permitindo assim a seleção e fusão de múltiplas informações de credibilidade e, consequentemente, o aprimoramento dos resultados da busca.

2 TRABALHOS RELACIONADOS

Nesta seção, apresentamos uma breve revisão sobre trabalhos envolvendo o uso de descritores de credibilidade na etapa de filtragem ou re-ranqueamento dos resultados de sistemas de recuperação da informação baseados em recuperação de imagens em mídias sociais.

Em [6], os autores propõem a exploração de dados de credibilidade dos usuários no processo de recuperação de imagens. O objetivo foi demonstrar um sistema de recuperação da informação capaz de retornar para o usuário resultados relevantes e diversificados. Para isso, na etapa de ranqueamento, os descritores de credibilidade foram utilizados para aumentar a relevância dos resultados considerando imagens provenientes de usuários com boas estimativas de credibilidade. Seus resultados mostraram que os

In: XV Workshop de Trabalhos de Iniciação Científica (WTIC 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pôsteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.

© 2018 SBC – Sociedade Brasileira de Computação.
ISBN 978-85-7669-435-9.

dados de credibilidade são eficazes quando utilizados em complemento com dados textuais e/ou visuais no processo de recuperação de imagens. Os escores de credibilidade utilizados foram obtidos comparando-se as tags dos usuários com os conceitos disponíveis na ImageNet [5], sem qualquer adaptação ao conjunto de dados de avaliação. Mesmo nesse cenário, os resultados foram considerados satisfatórios, ratificando o potencial dos descritores de credibilidade.

O trabalho em [8] apresenta um coleção de imagens projetada para dar suporte ao desenvolvimento de sistemas de recuperação que apresentem resultados de busca relevantes e diversificados. O conjunto de dados, coletados do Flickr¹ e Wikipedia², consiste em informações para 300 pontos turísticos em 35 países diferentes, e.g., museus, monumentos, edifícios, estradas, pontes, cavernas, entre outros. Para cada local, as seguintes informações são fornecidas: palavra-chave da localização (identificador textual exclusivo no conjunto de dados), número de localização (identificador numérico exclusivo) e coordenadas GPS (latitude e longitude em graus). Dentre os recursos presentes na base, há um conjunto de descritores de credibilidade, e.g.: o *faceProportion*, que indica a porcentagem de fotos de um usuário que apresentam faces humanas; o *tagSpecificity*, que é obtido calculando a especificidade média das tags de um usuário; e o *LocationSimilarity*, que dá ao usuário uma pontuação com base na similaridade entre as *tags* de localização das suas imagens e outras imagens que representam o mesmo local.

Em [3], dentre as abordagens apresentadas, é proposto o uso de um descritor de credibilidade para o ranqueamento dos resultados de buscas de imagens. Para isso, os autores utilizaram apenas o descritor *visualScore* para ordenar as imagens. Este descritor usa as tags das imagens do usuário como parâmetro de busca na coleção ImageNet [5]. Para cada tag do usuário idêntica a um conceito no ImageNet, é feito o comparativo entre as características visuais da imagem associada à tag e as características visuais das imagens recuperadas do ImageNet, obtendo um score de classificação para a tag. O score final é obtido através da média das scores obtidos para cada tag.

Para a combinação de múltiplos descritores de credibilidade, em [2], os autores propuseram uma abordagem multimodal com a utilização de metadados de imagens e informações de credibilidade do usuário. No arcabouço desenvolvido, os descritores de credibilidade são usados para ranqueamento de resultados de busca imagens em redes sociais. Os descritores foram integrados por meio de uma combinação linear simples, sendo os respectivos pesos fixados de forma empírica. Os resultados apresentaram melhor qualidade de ranking quando comparados com os originais obtidos pelo sistema de ranking da rede social, neste caso o Flickr.

3 MÉTODO PROPOSTO

O método aqui proposto tem como objetivo fornecer uma única informação de relevância por meio da combinação de múltiplos descritores de credibilidade. Esta informação é utilizada na etapa de re-ranqueamento dos resultados de uma determinada busca em um sistema de recuperação de imagens em rede social. Os descritores de credibilidade utilizados foram: *visualScore*, *faceProportion*,

tagSpecificity, *locationSimilarity*, *photoCount*, *uniqueTags*, *uploadFrequency*, *bulkProportion*, *meanPhotoViews*, *meanTitleWordCounts*, *meanTagsPerPhoto*, *meanTagRank* e *meanImageTagClarity* [8].

Para obter um score a partir da combinação dos descritores de credibilidade é utilizada a fusão linear (Equação 1), sendo d_i um elemento contido no conjunto dos descritores D , tal que $D = \{d_1, d_2, \dots, d_n\}$ e p_i o peso associado ao descritor d_i .

$$credScore = \frac{\sum_{i=1}^n p_i d_i}{\sum_{i=1}^n p_i} \quad (1)$$

A descoberta dos melhores pesos (p_i) para a combinação dos descritores foi realizada por meio da técnica de algoritmos genéticos, cuja estrutura geral é apresentada na Figura 1. No AG implementado, a solução (indivíduo no AG) foi modelada como um vetor de pesos (genótipo) formado por valores de ponto flutuante (genes), em que cada valor representa o peso de um descritor.

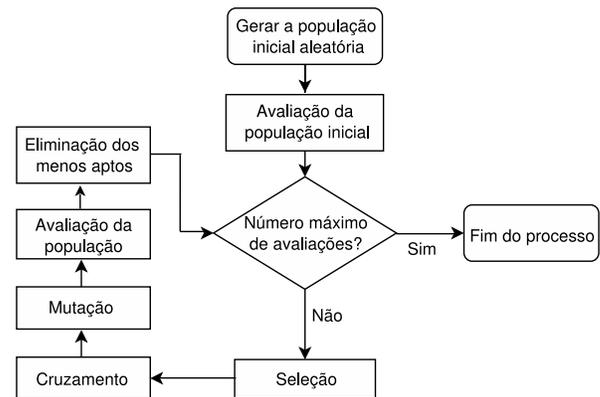


Figura 1: Fluxograma do algoritmo genético implementado.

A primeira etapa do algoritmo é a criação uma população aleatória de indivíduos, ou seja, vetores cujos pesos são inicializados com valores aleatórios. Em seguida, esses indivíduos são avaliados e submetidos ao operador de seleção por torneio binário. Feito isso, os indivíduos selecionados são submetidos à etapa de reprodução, onde são submetidos aos operadores de cruzamento e mutação. Finalizado o ciclo de reprodução, a nova população é avaliada e os indivíduos menos aptos são descartados. Este processo é repetido até atingir o número máximo de avaliações.

A avaliação da aptidão (*fitness*) de cada indivíduo é feita da seguinte forma: dado um ranking de entrada, a fusão linear dos descritores de credibilidade é realizada utilizando os pesos fornecidos pelo indivíduo avaliado. Em seguida, as imagens presentes no resultado original são re-ranqueadas utilizando o score de credibilidade obtido. Uma medida de avaliação é utilizada para avaliar a eficácia do novo ranking. O valor obtido com a medida de avaliação é utilizado como a aptidão do indivíduo.

Após a realização de todas as etapas do AG, o melhor indivíduo da última geração reproduzida é definido como a melhor configuração de pesos para os descritores.

¹<https://www.flickr.com>

²<https://www.wikipedia.org>

4 METODOLOGIA EXPERIMENTAL

O AG desenvolvido tem como base o modelo NSGA-II. Para o auxiliar no desenvolvimento, foi utilizado o framework JMetal³. O comportamento do AG depende da configuração de alguns parâmetros. A Tabela 1 apresenta a configuração utilizada no AG implementado. Os parâmetros de probabilidade de cruzamento e de mutação definem as chances que cada indivíduo tem de ser submetido a estes operadores, enquanto o tamanho da população e a quantidade de sobreviventes definem, respectivamente, a quantidade de indivíduos em uma população e o número de indivíduos que serão replicados para a próxima geração. A quantidade de avaliações define o número máximo de indivíduos que devem ser avaliados.

Cruzamento	Mutação	População	Sobreviventes	Avaliações
90%	7%	200	10	5000

Tabela 1: Parâmetros utilizados no algoritmo genético.

Para determinar a probabilidade de cruzamento e mutação do AG foram utilizados os valores padrões para o modelo NSGA-II implementado no framework JMetal.

O AG precisa passar por uma etapa de busca de soluções (o que garante a evolução dos indivíduos) e por uma etapa de testes (para avaliar quão boa foi a solução final). Assim, de modo a avaliar adequadamente a capacidade do algoritmo de encontrar boas soluções e que sejam eficazes para novos dados (consultas), foi definida a utilização de conjuntos de dados distintos para cada etapa. Com base nesta ideia e com o objetivo de ter um grande conjunto de dados, optou-se por unir os conjuntos de desenvolvimento e testes apresentados em [8], totalizando 222 consultas com cerca de 300 imagens no resultado cada uma. Com a unificação dos conjuntos de dados, optamos por aplicar a técnica da validação cruzada. A validação cruzada ajuda a avaliar a capacidade de generalização do AG, ou seja, a possibilidade de estimar a eficácia da solução encontrada pelo AG quando submetida a um novo conjunto de dados.

O modelo de validação cruzada utilizado foi o *k-fold cross validation*. Esta técnica particiona aleatoriamente o conjunto de dados em *k* subconjuntos mutuamente exclusivos de tamanhos iguais. Para o treinamento são utilizados *k* - 1 subconjuntos, enquanto o conjunto restante é utilizado para testes. Este processo é realizado *k* vezes alternando de forma circular o subconjunto de testes. Para os experimentos realizados foi definido *k* = 5. Desse modo, em cada etapa da validação cruzada aproximadamente 178 consultas foram utilizadas para descoberta dos pesos e 44 consultas foram utilizadas para avaliação.

Como função de adequação dos indivíduos gerados durante a etapa de treinamento do AG, foram avaliadas as medidas *Mean Average Precision* (MAP) e *Precision* (P@N), especificamente P@10, P@20 e P@50. A medida P@N indica a proporção de imagens relevantes dentre as N primeiras imagens retornadas (Equação 2). Neste trabalho, a função *Average Precision* (AP) foi computada como a média dos valores de *Precision* da curva interpolada de *Precision-Recall* de 11 pontos. A medida MAP é então computada como a média dos valores de AP considerando múltiplas consultas.

$$P@N = \frac{\{Itens\ relevantes\} \cap \{Itens\ retornados\}}{\{Itens\ retornados\}} \quad (2)$$

Utilizando a mesma configuração de validação cruzada, os experimentos foram executados separadamente para cada uma destas funções de *fitness*.

Por fim, a etapa de teste foi realizada, na qual os pesos encontrados pelo AG foram utilizados na Equação 1 para obter um score final de credibilidade. Este score é utilizado para re-ranquear os resultados para as consultas presentes no subconjunto de teste. O novo ranking gerado foi então avaliado utilizando a medida MAP e curvas de P@N. Os resultados são apresentados até a profundidade máxima de 50, cobrindo aquilo que, de modo geral, é apresentado nas primeiras páginas dos sistemas de busca.

5 RESULTADOS E DISCUSSÕES

Os rankings obtidos com o método aqui proposto foram comparados com os rankings originais obtidos do Flickr, com o método proposto em [3] e com a utilização dos melhores pesos obtidos com o método em [2]. Primeiro foi feito um comparativo entre os resultados obtidos utilizando as medidas P@10, P@20, P@50 e MAP como função de *fitness*. Os resultados em termos de P@N são apresentados na Figura 2.

Os resultados do método proposto com avaliação de *fitness* baseada na medida P@10 sugerem que, como esta medida considera apenas o topo do ranking (as 10 primeiras posições), não foi sensível o suficiente para capturar variações no ranking e permitir ao AG avaliar adequadamente as soluções encontradas. Analisando-se as curvas para as medidas P@20 e P@50, percebe-se que com o aumento da profundidade de análise do ranking, melhores resultados são obtidos. Por fim, analisando-se a curva do resultado obtido com avaliação de *fitness* baseada na medida MAP, percebe-se uma degradação do resultado. Estima-se que, como a medida MAP leva em consideração todo o ranking, pequenas variações de posições na parte superior do ranking tem menor impacto no resultado computado, tornando a avaliação menos sensível a estas variações.

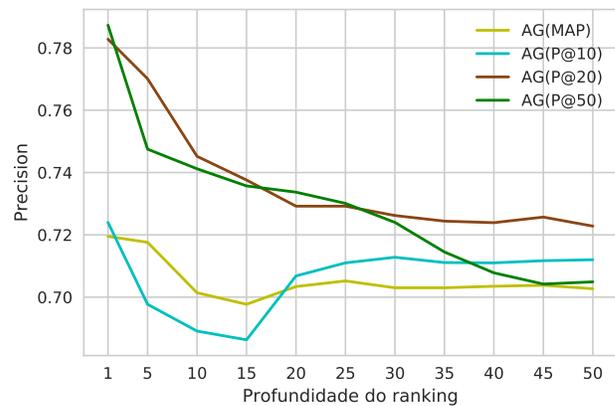


Figura 2: Comparativo dos resultados obtidos com o AG.

³<http://jmetal.sourceforge.net>

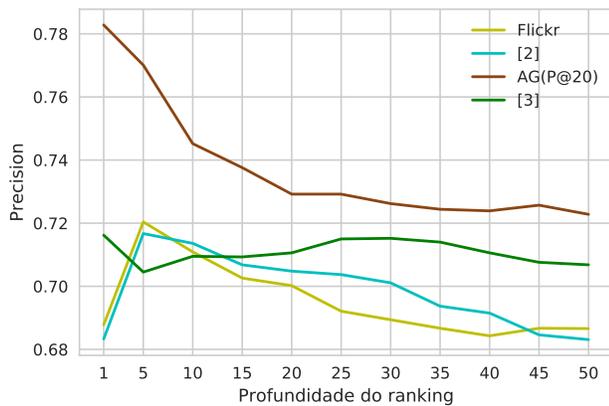


Figura 3: Comparativo do AG(P@20) com baselines.

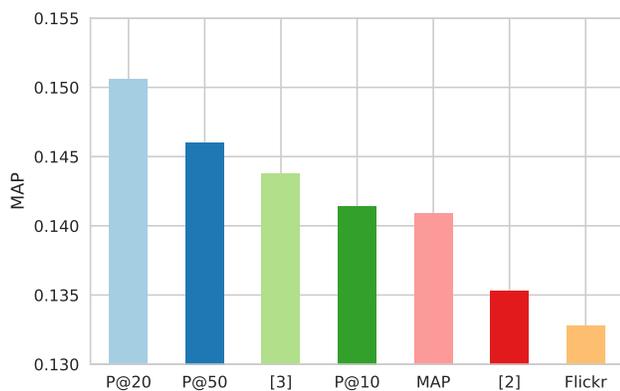


Figura 4: Avaliação comparativa dos resultados utilizando a medida MAP.

Por meio da Figura 2, percebe-se que, dentre as funções de *fitness* utilizadas, o melhor resultado foi obtido pela P@20, seguido de P@50. Estima-se que a utilização da medida P@20 teve vantagem em relação à P@50 por analisar uma parte do ranking que tende a ter uma maior proporção de itens relevantes em relação à quantidade de itens recuperados, ao mesmo tempo que avalia dados suficientes para ajudar o AG à encontrar uma solução adequada. Estes resultados são comparados com os obtidos utilizando os melhores pesos encontrados em [2], com os resultados obtidos utilizando o método proposto em [3] e com o ranking original do Flickr. A Figura 3 apresenta o comparativo entre os *baselines* e o resultado obtido utilizando P@20 como *fitness*. De forma geral, pode-se perceber que houve ganho expressivo em relação aos *baselines*.

A Figura 4 apresenta os resultados obtidos com as quatro medidas de *fitness* em comparação aos *baselines*, utilizando MAP como medida de avaliação. Dentre as medidas de aptidão utilizadas, mais

uma vez pode-se destacar os resultados obtidos utilizando P@20 e P@50. Ambas conseguiram resultados superiores em comparação aos *baselines*.

6 CONCLUSÕES E TRABALHOS FUTUROS

Neste trabalho, apresentamos um modelo de re-ranqueamento dos resultados de um sistema de recuperação de imagens em redes sociais, utilizando a combinação de descritores de credibilidade através da fusão linear, cujos pesos são definidos por um algoritmo genético. Os resultados preliminares apresentam-se superiores aos *rankings* originais do sistema de busca da rede social e ao melhor *baseline*. Além disso, pode-se observar que a escolha da função de aptidão tem impacto direto na descoberta dos pesos e consequentemente na eficácia do método proposto.

Vale destacar que apenas uma configuração paramétrica foi avaliada para as etapas do algoritmo genético e portanto uma nova etapa de experimentação será realizada para melhor avaliar o método proposto. O ajuste destes parâmetros poderá permitir a obtenção de melhores soluções provenientes do AG, pois sintonizando seus parâmetros, o AG tende a se adequar às particularidades do problema. Outra etapa de grande importância é a experimentação de novos operadores para seleção, mutação e cruzamento, que podem influenciar de forma direta o processo de busca dos melhores pesos para os descritores de credibilidade.

O método proposto foi avaliado apenas em termos do aprimoramento do ranking, mas seus resultados podem ser utilizados para potencialização de métodos de diversificação visual que naturalmente são beneficiados por rankings de maior qualidade. De modo adaptativo, o método proposto pode ainda ser utilizado para descoberta de soluções para outros cenários de busca, diferentes bases de dados, ou tipos de objetos digitais (textos, vídeos, etc.).

REFERÊNCIAS

- [1] Ricardo Baeza-Yates and Berthier Ribeiro-Neto. 2011. *Modern Information Retrieval - the concepts and technology behind search, Second edition*. Pearson Education Ltd., Harlow, England.
- [2] Rodrigo Tripodi Calumby, Iago B. A. do C. Araujo, Vinicius Pereira Santana, Javier A. V. Muñoz, Otávio Augusto Bizetto Penatti, Lin Tzy Li, Jurandy Almeida, Giovanni Chiachia, Marcos André Gonçalves, and Ricardo da Silva Torres. 2015. Recod @ MediaEval 2015: Diverse Social Images Retrieval. In *Working Notes Proceedings of the MediaEval Workshop*.
- [3] Xi Chen, Haokun Liu, Zhi-Hong Deng, and Yunlun Yang. 2015. Retrieving Relevant and Diverse Image from Social Media Images.. In *Working Notes Proceedings of the MediaEval Workshop*.
- [4] Ritendra Datta, Dhiraj Joshi, Jia Li, and James Ze Wang. 2008. Image retrieval: Ideas, influences, and trends of the new age. *Comput. Surveys* 40, 2 (2008).
- [5] J. Deng, W. Dong, R. Socher, L. Li, Kai Li, and Li Fei-Fei. 2009. ImageNet: A large-scale hierarchical image database. In *IEEE Conference on Computer Vision and Pattern Recognition*. 248–255.
- [6] Alexandru Lucian Ginsca, Adrian Popescu, Bogdan Ionescu, Anil Armagan, and Ioannis Kanellos. 2014. Toward an estimation of user tagging credibility for social image retrieval. In *Proceedings of the 22nd ACM International Conference on Multimedia*. ACM, 1021–1024.
- [7] Bogdan Ionescu, Alexandru-Lucian Ginsca, Bogdan Boteanu, Adrian Popescu, Mihai Lupu, and Henning Müller. 2014. Retrieving Diverse Social Images at MediaEval 2014: Challenge, Dataset and Evaluation.. In *Working Notes Proceedings of the MediaEval Workshop*.
- [8] Bogdan Ionescu, Adrian Popescu, Mihai Lupu, Alexandru Lucian Ginscă, Bogdan Boteanu, and Henning Müller. 2015. Div150cred: A social image retrieval result diversification with user tagging credibility dataset. In *Proceedings of the 6th ACM Multimedia Systems Conference*. ACM, 207–212.

Using images to extend smart object discovery in an Internet of Things scenario

Romulo Silva

Virtual University Institute
Federal University of Ceara
60451-970, Fortaleza, Ceará, Brazil
romulosn@alu.ufc.br

Windson Viana

Group of Computer Networks,
Software Engineering and Systems
Federal University of Ceara
60451-970, Fortaleza, Ceará, Brazil
windson@great.ufc.br

Paulo Filipe Dantas

Group of Computer Networks,
Software Engineering and Systems
Federal University of Ceara
60451-970, Fortaleza, Ceará, Brazil
paulodantas@great.ufc.br

ABSTRACT

The Internet of Things (IoT) represents a new paradigm in the Internet history and in the way people interact with everyday objects. Academics and several industry segments have been working over the past years to make this vision possible. It is estimated that there will be dozens or hundreds of devices simultaneously connected to the user's network (e.g., in their home) in the next years, which can make the discovery and interaction with smart object more complicated to IoT users increasingly. This work proposes a smart objects discovery approach using image recognition, which aims to make this task quicker and more selective from an user perspective. An initial assessment has shown that the proposed mechanism can reduce the discovery time in a scenario with several devices, and additionally ensure a good level of user satisfaction.

KEYWORDS

IoT, Smart object discovery, Pointing, image recognition

1 INTRODUCTION

The Internet emerged in the 60s with the primary goal of sharing information between large computers. Subsequently, the Internet connected not only machines but also people. Several technological solutions have been developed to simplify people's lives and computerize everyday tasks. Nowadays, the Internet is undergoing another paradigm shift, in which not only machines and human beings are connected, but everyday objects as well. [7] This new scenario is known as Internet of Things (IoT). Atzori et al [1] define IoT as the constant and invisible presence of a huge variety of computing devices in people's daily lives, which are capable of interacting with each other in order to cooperatively perform certain tasks. Those computing devices, which may vary terms of functionality and resources, are commonly called smart objects (SO). As examples of SOs we can mention smart TVs, Wi-Fi lamps, and connected thermostats.

Due to their growing popularity and easy market access, smartphones have become main way by which IoT users discover, connect to, and interact with SO's [5]. Smartphones can be seen as universal controls or environment browsers that work actively on the discovery of devices of interest to the user [4].

The most common smart object discovery approaches aim to discover and make all SO's within a network available to users. However, those approaches can become problematic when there are several SO's within reach. The Cisco Internet Business Solutions Group (Cisco IBSG) estimates that there will be 50 billion smart objects connected by 2020. The number of objects connected to the network can reach a mark of 6.58 per person. [3].

Under those circumstances, a more selective and appropriate smart object discovery approach is necessary to reduce the interest options of SO's in a network. The use of digital cameras – tools already available today on smartphones – can filter the options in the process of discovering and interacting with SO's. This work, therefore, aims to develop a smart object discovery approach using image recognition techniques. In order to achieve the main goal, the following specific goals were established: (i) analyze methods of interaction with smart objects currently used by mobile devices; (ii) identify smart object discovering challenges related to each method; (iii) evaluate a smart object discovery solution using smartphone cameras combined with a network-based discovery technique. Thus, this research is centered on the following question: how to cause the smart object discovery to be more selective and appropriate to the context of an IoT user?

2 INTERNET OF THINGS

The Internet of Things represents a vision in which the Internet extends to the real world and reaches everyday objects. These objects, now called smart objects, are capable of interacting and communicating with other SO's, users, and the environment in which they live. Thus, SO' are able to provide services and applications, autonomously and cooperatively [6]. IoT makes computing truly ubiquitous [10].

An important part of IoT systems is the discovery and configuration of SOs available to the users. In fact, SO discovery is a similar problem to the resource or the discovery challenge.

Service discovery is a known term from the distributed systems field. It is defined as the process of finding suitable services to perform a specific task [2]. These services may be presented in the form of available software abstractions – such as an e-mail server – or in the form of computing resources or devices (e.g., printers on a network). Dynamic Host Configuration Protocol (DHCP) and Simple Service Discovery Protocol (SSDP), Universal Plug and Play (UPnP), are examples of protocols that have service discovery mechanisms. Each service discovery protocol has characteristics that make it more suitable to a given domain.

In: XV Workshop de Trabalhos de Iniciação Científica (WTIC 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pósteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.

© 2018 SBC – Sociedade Brasileira de Computação.
ISBN 978-85-7669-435-9.

2.1 UPnP

In our research we adopt UPnP as a starting point. UPnP¹ is a set of protocols that provide data description, easy configuration, automatic discovery of services and control networked devices. The UPnP architecture includes a combination of protocols such as HTTP (Hypertext Transfer Protocol), SOAP (Simple Access Object), XML (Extensible Markup Language) and SSDP (Simple Service Discovery Protocol), which support integrated and seamlessly device discovery, connection, control and data exchange.

UPnP has two general classifications for devices: controlled devices and control points. A controlled device acts as a server, offering services and responding to control points requests. A control point, in the other hand, is a device that uses UPnP to consume services and control devices.

2.2 Physical Mobile Interactions

Other important aspect of our research is the user interaction with SOs. We follow some insights proposed by Enrico Rukzio et al in [9]. In that work, the authors proposed, evaluated and compared four types of physical mobile interactions using mobile devices: touch, pointing, scanning and user-mediated.

Touching is the interaction that involves an action where the user's device is put in physical contact with a device that it wants to interact with. In this type of interaction, the user needs to know the location of the object and to be aware of the possibility of performing this sort of interaction with the device. Examples of this type of interaction can be achieved using technologies such as Near Field Communication (NFC).

Pointing allows the user to interact with object by pointing to it with a mobile device. Pointing has been widely used in augmented reality applications through the mobile device camera. Applications that use pointing as a form of interaction perform visual recognition of bar codes, QR code, reference markers or patterns.

Scanning provides information from a geographical location or by connecting to a network, such as a list of available SO's in a given perimeter. One of the advantages of this approach is the possibility of interacting with SO's with no prior knowledge of their availability [9].

Finally, the user-mediated interaction is characterized by the user's action of entering some information previously provided by the SO in order to establish a connection with it, such as URLs printed in SO's that provide access to the object information and functions.

3 PROPOSAL

The main purpose of this work is to integrate and combine the pointing interaction with network based scanning mechanisms. As a result of doing so, we expect to improve smart object discovery services (e.g. UPnP) by using image recognition techniques.

3.1 Proof of Concept

In order to illustrate the smart object discovery approach proposed by this work, a Proof of Concept (PoC) called SmartUPnP (Figure 1) was developed. This PoC consists of an Android² application,

¹UPnP: <https://openconnectivity.org/developer/specifications/upnp-resources>

²Android: <https://developer.android.com/about/>

written in Java³, which acts as a UPnP control point. It is capable of discovering devices within a network by using scanning and pointing physical interactions. TensorFlow⁴ was used to add image classification features to SmartUPnP. TensorFlow uses neural convolutional networks to classify images, group them by similarity and identify objects within a given scene.

In order to be able to classify objects, 4809 public domain images from 29 different objects were collected. Those images were then used as inputs to create a data model with Tensorflow's MobileNets⁵ architecture.

To simulate a real usage scenario of the PoC, a second Java application was also developed to simulates several smart objects that would be commonly found in smart homes, such as a refrigerator, a television and a vacuum cleaner. This simulating application also used UPnP to create those 17 devices, publish their services and receive commands sent by control points.

To implement the UPnP architecture on SmartUPnP, it was used Cling⁶, which is an open source library – written in Java – with the goal of providing a programming interface compatible with the UPnP architecture layers.

When identifying an object through the camera, the SmartUPnP app searches for devices within reach that may correspond to the identified object name.

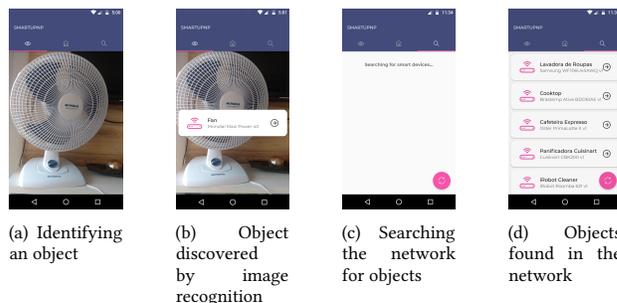


Figure 1: SmartUPnP interface

4 EVALUATION

In the presented context, the smart object discovery approaches using pointing and scanning techniques were evaluated by users. The experiment occurred in a residential setting, simulating a real use case.

4.1 Methodology

The experiment was carried out with a total of 9 participants, among university students and technology professionals. Although 9 people is a small number, researchers such as Nielsen [8] demonstrate that it is possible to find the majority of problems on a software product with only 5 users.

³Java: <https://www.java.com/about/>

⁴TensorFlow: <https://www.tensorflow.org/>

⁵MobileNets: <https://ai.googleblog.com/2017/06/mobilenets-open-source-models-for.html>

⁶Cling: <https://github.com/4thline/cling/>

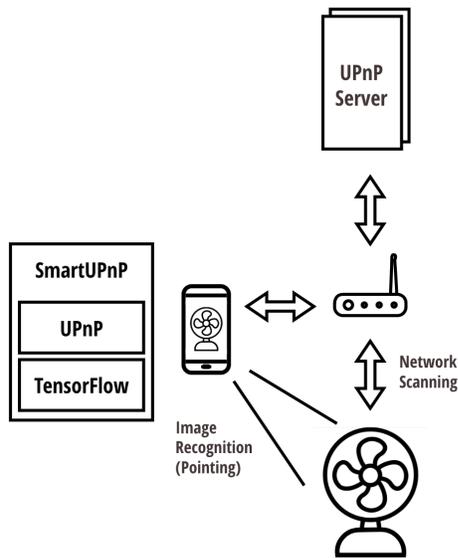


Figure 2: SmartUPnP Architecture

Four images of possible intelligent objects were fixed in a room: a table lamp, a fan, a refrigerator and a vacuum cleaner. All users were familiar with the use of smartphones. Firstly, a pre-test questionnaire was applied in order to identify the participants' prior perception of how interaction between smartphones and real-world objects could be performed.

Next, the experiment scenario was presented to each participant: a smart home with several devices connected to the Internet. Each participant was then instructed to find each of the four objects using the network search method and then the image recognition approach. The entire experiment was timed in order to collect the time taken by each user to discover each object using each of the two types of discovery approach.

After the test, another questionnaire was applied in order to evaluate the participants' experience with each of the discovery methods.

4.2 Results

4.2.1 User profile

All participants in the experiment were smartphone users and already had some knowledge about the term "smart object". 6 of them reported that they currently use or have used a SO before. All participants related SO's to devices that are able connect to the Internet or perform tasks autonomously.

When asked about their opinion on the use of smartphones to interact with SO's, 8 participants considered it very useful or useful to use a smartphone to access information about the state of an object, such as the temperature of the oven or if the room lights are on (Figure 3). When asked how useful they considered using a smartphone to control an SO, such as adjusting the brightness of an environment or locking the doors of a car, all participants considered it very useful or useful.

Participants' views on when they would use their smartphones were also requested. 6 participants answered that they would use their smartphones only when they were away from the object, while only 3 said they would always use them.

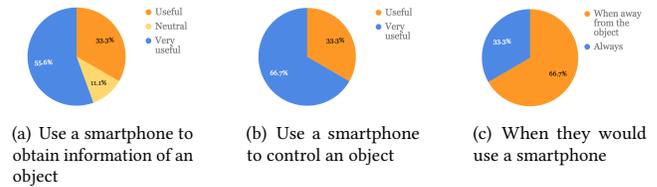


Figure 3: Use of smartphones to interact with objects

Before beginning the experiment, participants were asked to indicate their preferences as to the type of interaction they would find most interesting in an IoT context (Figure 4). Only 6 claimed to be interested in using the camera to point to an object, while 8 would choose a network search approach.

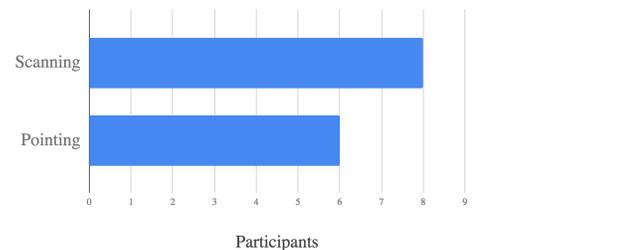


Figure 4: Participants' preferences as to the type of interaction

4.2.2 Experiments

During the experiment, all participants had some difficulties with both discovery approaches. However, after the test, they all said they were generally satisfied with the way the objects were discovered.

Using the network scanning approach, users took considerable time to locate at least one of the objects. Search times were higher usually for the first objects of the experiment, reaching 31.13 seconds. This was probably due to the fact that users were still familiarizing themselves with the application interface and the list of available connected objects. The discovery mean time using the network scanning approach was 9.51 seconds [7.53-11.49s], meanwhile the use of image recognition had an average time of 6.53 seconds [5, 22-7.84s]. That shows evidence that the image recognition discovery method can be faster than the network scanning method in a scenario with several connected devices.

On a scale of 1 to 5, where 1 represents not satisfied and 5 represents very satisfied, the network scanning method obtained an average of 3.6 points (Figure 6). Participants argued that they found this type of discovery more suitable to use in a context with few connected objects. They would also like that the application interface provided some options for filtering the available objects. This

filtering could be done by name or device type, or even by usage context, such as the user location and time of day.

On a scale of 1 to 5, where 1 is not satisfied and 5 is very satisfied, the image recognition discovery approach has averaged 4.3 points (Figure 6). Despite the long response time in some cases, and the fact that the application would sometimes mistakenly show more than one object, the participants stated that they found this approach to be faster, simpler and more natural. However, they also commented that they would like the application to be more precise in terms of identification.

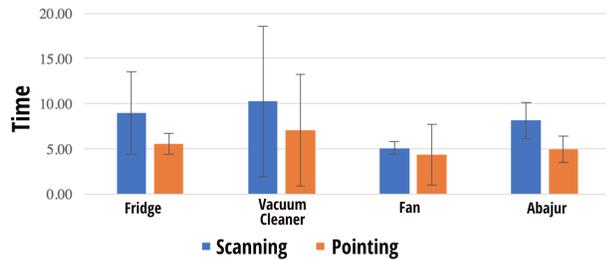


Figure 5: Smart objects discovery times

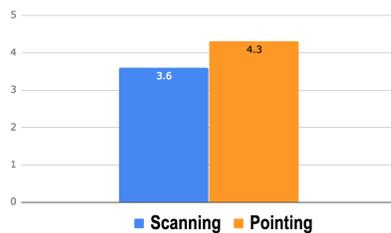


Figure 6: User satisfaction score

Participants were generally satisfied with this work proposal. From the positive comments and points of satisfaction obtained, it is possible to validate the smart objects discovery by image recognition in an IoT context. The image recognition method is as satisfactory as the network scanning method according to the users and the time performance test. The difficulties and issues found during the experiment can be addressed by improving the usability of the graphical user interface and improving the image recognition precision.

5 FINAL CONSIDERATIONS

With the rise of the Internet of Things, it is expected that there will be dozens of devices connected to the same network. In this context, several aspects related to the discovery of smart objects are still to be addressed. This work presents an smart object discovery approach that uses image recognition in order to make the discovery processes faster and more selective from the point of view of an IoT user.

A proof of concept has been developed using technologies already known and used by academia and industry (i.e TensorFlow

and UPnP). This solution was based on two types of physical mobile interaction proposed by [9], scanning and pointing.

An experiment evaluated the users' perception and the effectiveness of a smart object discovery approach using image recognition. The results showed that the proposal is feasible and can be used in a real context. It was also possible to conclude that users were satisfied with the way objects were discovered using a smartphone camera. In most cases tested, in a scenario with many connected devices, users were able to discover a connected object using the camera of their smartphones more quickly when compared to a network scanning approach.

During the research, some future works were identified that could improve the proposal of this work. The limited integration between the TensorFlow image recognition and the UPnP discovery mechanism showed the need to create a software component that would more specifically support smart object discovery by image recognition.

It would also be interesting to conduct new experiments with different machine learning algorithms to find out if other approaches or set of approaches could minimize time and error while detecting objects in a scene. Image recognition algorithms are often computationally expensive and, consequently, consume mobile devices energy resources quickly. Therefore, it would also be relevant to investigate the energy efficiency of the this work and what other approaches could offer a better cost-benefit in terms of energy.

REFERENCES

- [1] Luigi Atzori, Antonio Iera, and Giacomo Morabito. 2010. The internet of things: A survey. *Computer networks* 54, 15 (2010), 2787–2805.
- [2] Marco Crasso, Alejandro Zunino, and Marcelo Campo. 2008. Easy web service discovery: A query-by-example approach. *Science of Computer Programming* 71, 2 (2008), 144–164.
- [3] Evans Dave. 2011. The internet of things: how the next evolution of the internet is changing everything. *CISCO white paper* (2011).
- [4] Nayeem Islam and Roy Want. 2014. Smartphones: Past, present, and future. *IEEE Pervasive Computing* 13, 4 (2014), 89–92.
- [5] Feida Lin and Weiguo Ye. 2009. Operating system battle in the ecosystem of smartphone industry. In *Information Engineering and Electronic Commerce, 2009. IEEEC'09. International Symposium on*. IEEE, 617–621.
- [6] Friedemann Mattern and Christian Floerkemeier. 2010. From the Internet of Computers to the Internet of Things. In *From active data management to event-based systems and more*. Springer, 242–259.
- [7] NIC.br. 2014. A Internet das Coisas, explicada pelo NIC.br. (July 2014). <https://www.youtube.com/watch?v=jlkvzcG1UMk>
- [8] Jakob Nielsen and Thomas K. Landauer. 1993. A Mathematical Model of the Finding of Usability Problems. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems (CHI '93)*. ACM, New York, NY, USA, 206–213. <https://doi.org/10.1145/169059.169166>
- [9] Enrico Rukzio, Gregor Broll, Karin Leichtenstern, and Albrecht Schmidt. 2007. Mobile interaction with the real world: An evaluation and comparison of physical mobile interaction techniques. In *European Conference on Ambient Intelligence*. Springer, 1–18.
- [10] Mark Weiser. 1999. The computer for the 21st century. *Mobile Computing and Communications Review* 3, 3 (1999), 3–11.

EasyContext: Facilitando o desenvolvimento de aplicações com a Awareness API

Davi Batista Tabosa
Universidade Federal do Ceará
Av. Mister Hull, s/n
Fortaleza, Ceará 60451-970
davitabosa12@gmail.com

Windson Viana
Universidade Federal do Ceará
Av. Mister Hull, s/n
Fortaleza, Ceará 60451-970
windson@virtual.ufc.br

ABSTRACT

Due to the popularity of smartphones and their built-in sensors, Context-Aware Mobile (CAM) applications are on the rise. Some libraries and APIs help developers to create CAM applications, such as the Google Awareness API. However, the initial configuration, the write of context-aware rules, and code readability persist being complex tasks, specially for beginner developers. This paper proposes an approach to ease the development of CAM applications using Awareness API. Our proposal, called EasyContext, includes a Web Configurator and an Android Framework, which hides code complexity and the Awareness API configuration. A PoC using the EasyContext is also presented.

KEYWORDS

Context-Awareness, Google Awareness API, Ubiquitous Computing

1 INTRODUÇÃO

Aplicações móveis sensíveis ao contexto (e.g., WAZE) e jogos móveis baseados em informações contextuais dos jogadores (e.g., Pokemon Go) tem tido ascensão notória nos últimos anos devido à popularização dos dispositivos móveis. Embutido nestes aparelhos, há vários sensores como giroscópios, acelerômetros, bússolas e GPS[10]. Este hardware nativo combinado com algoritmos de monitoramento e reconhecimento de situações permitem que aplicações e jogos móveis se adaptem conforme à situação do ambiente ou do usuário que as acessa. Essas aplicações, ditas CAM (*Context-Aware and Mobile* apps, ou Móveis e Sensíveis ao Contexto) executam a maior parte de sua interação com o usuário em um dispositivo móvel e adaptam o seu comportamento (e.g., conteúdo, interface do usuário) para o contexto atual do usuário e do ambiente [1][10]. O desenvolvimento de aplicações CAM tem sua gama de desafios, dentre os quais, podem ser citados: desafios inerentes à heterogeneidade dos dispositivos e plataformas, à complexidade do código voltado ao acesso de sensores e à dificuldade da definição de regras contextuais (e.g., comportamento que o sistema deve ter quando o usuário entrar em uma determinada região) [9][7][3][4].

Durante o desenvolvimento de um software é comum o programador encontrar estas situações problemáticas ao estruturar o projeto, sendo muitas delas recorrentes. Muitas pesquisas e ferramentas comerciais surgiram, então, com o objetivo de dirimir

esses desafios, como a utilização de *frameworks* e/ou plataformas de *middleware* para uniformizar a sintaxe da aquisição e gerenciamento de informações contextuais[6][5][8]. Seguindo essa vertente, a Google desenvolveu uma API focada na aquisição de contexto na plataforma Android, a Google Awareness API¹. Essa API permite ao desenvolvedor a aquisição de sete grupos de informações contextuais e a criação de regras contextuais que as combinem por meio do Google Play Services. Assim, para dispositivos da plataformas Android, a Awareness API oferece uma solução que não precisa da prévia instalação de outro software (como um *framework* ou *middleware*). No entanto, algumas funcionalidades do Google Awareness API, como a criação de regras contextuais, necessitam ainda de um processo complexo e verboso para serem implementadas. Dentro deste contexto, este trabalho propõe uma ferramenta de configuração da biblioteca do Google Awareness aliado com o uso de *framework* de desenvolvimento criado com práticas de padrões de projetos. A abordagem, chamada de EasyContext, visa facilitar o desenvolvimento de aplicações sensíveis ao contexto, em especial, para programadores iniciantes neste tipo de domínio. O trabalho mais relacionado a nossa abordagem é a Ferramenta CRITICAL [3], que dispõe de um editor visual para a criação de regras contextuais. A partir do modelo, é gerado um projeto Android com as regras contextuais já configuradas, prontas para serem utilizadas no projeto. Contudo a CRITICAL utiliza o LoCCAM [7] para a aquisição de informações contextuais que possui poucas informações contextuais implementadas. A seguir, detalha-se a fundamentação teórica deste trabalho e sua proposta principal.

2 SENSIBILIDADE AO CONTEXTO

Segundo Dey [2], "Contexto é qualquer informação que pode ser utilizada para caracterizar a situação de uma entidade". Esta entidade pode assumir várias formas, como, "uma pessoa, lugar, ou objeto considerado relevante para a interação entre um usuário e uma aplicação, incluindo o usuário e a aplicação em si". Uma aplicação sensível ao contexto é definida como "aquela que utiliza contexto para prover informações relevantes e/ou serviços para o usuário, onde a relevância depende da tarefa do usuário". Uma das vantagens de utilizar sensibilidade ao contexto em aplicações móveis é a adaptação de seu comportamento dependendo do contexto onde o usuário está inserido (e.g., localização, temperatura), promovendo uma interação mais natural[3].

Um dos desafios pertinentes no desenvolvimento de aplicativos sensíveis ao contexto é a heterogeneidade de plataformas (e.g., sensores, dispositivos de acesso) [10]. Por exemplo, no mesmo ecossistema Android, existem vários tipos de smartphones, cada um com

In: XV Workshop de Trabalhos de Iniciação Científica (WTIC 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pósteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.
© 2018 SBC – Sociedade Brasileira de Computação.
ISBN 978-85-7669-435-9.

¹<https://developers.google.com/awareness/>

hardware distinto. Certos modelos podem possuir ou não sensores que podem interferir na precisão da inferência de contexto. Além da própria complexidade inerente da programação deste tipo de aplicação. Uma estratégia para dirimir estes problemas é a criação de APIs e plataformas de middleware, como a Google Awareness API, que introduz uma camada de acesso ao hardware de forma uniforme, tornando transparente a aquisição de contexto além de realizar inferência a partir de dados brutos.

A Google Awareness API foi lançada em 2016, exclusivamente para a plataforma Android como parte da suite de serviços do Google Play Services². Ela é dividida em duas APIs, chamadas Snapshot e Fence. A API Snapshot obtém os dados de sensores e realiza a inferência no momento em que seu método é executado. Já a API Fence representa uma regra de contexto[10], em que um trecho de código é executado caso esta regra torne-se verdadeira ou falsa. Para o trecho de código ser executado, ele deve estar configurado e registrado utilizando o componente Broadcasts³ da plataforma Android. Esse tipo de abordagem é contra intuitivo, pois requer que desenvolvedor realize uma série de etapas que aumentam a complexidade de codificação.

A Awareness API oferece sete categorias de informações contextuais: *time* (hora local), *location* (coordenadas geográficas), *place* (lugar), *activity* (detecção de atividades do usuário), *beacons* (detecção de beacons próximos), *headphones* (detecção da presença de fones de ouvido) e *weather* (condições do clima). Algumas categorias que estão presentes na API Snapshot, como *weather* e *place*, não estão presentes na API Fence, o que impossibilita a criação de regras contextuais para essas categorias.

3 O FRAMEWORK EASYCONTEXT

Este trabalho propõe a criação do EasyContext, um *framework* para auxiliar desenvolvedores a inserir sensibilidade ao contexto em suas aplicações Android de forma menos verbosa e mais simples. Ele funciona como uma extensão do Google Awareness API. A abordagem é dividida em duas partes: um configurador web para a escrita de regras contextuais e uma biblioteca Java para a interpretação destas regras, além da gerência de registro de Fences e das abstrações para a invocação de métodos da Awareness API.

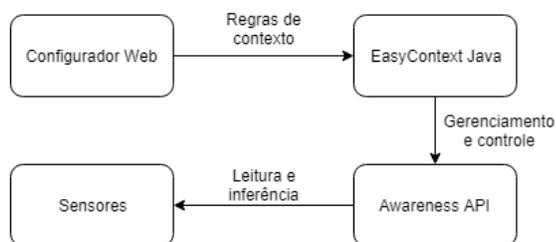


Figura 1: Arquitetura simplificada do EasyContext.

No EasyContext, uma regra contextual possui duas partes: Condição contextual e Ação. Uma condição contextual é uma

²<https://developer.android.com/distribute/play-services/>

³<https://developer.android.com/guide/components/broadcasts>

cláusula ou conjunto de cláusulas que definem um contexto. Ação é um conjunto de instruções que serão executadas caso uma condição contextual ocorra. Um exemplo de regra conceitual seria: "Caso o usuário entre em um hospital, configure o celular em modo silencioso". Neste caso, a condição contextual seria "entrar em um hospital" e o fenômeno resultante caso esta condição seja verdadeira seria "configurar o celular em modo silencioso".

3.1 Configurador Web

Algumas ferramentas de configuração de regras de contexto são feitas como plugins de IDEs. Isso pode ser um problema caso a IDE em questão entre em desuso. Por exemplo, a Ferramenta CRITiCAL[3] foi desenvolvida como plugin para a IDE Eclipse. Porém, atualmente a IDE padrão para desenvolvimento Android deixou de ser o Eclipse e passou a ser o Android Studio, o que deixou a CRITiCAL obsoleta sem um grande retrabalho. Para evitar esta situação, o configurador do EasyContext foi feito como um aplicativo Web, deixando de ser dependente de IDEs. O artefato final gerado é um documento no formato JSON para ser incorporado no projeto Android do desenvolvedor. Na Figura 2, é exposto um exemplo de documento gerado com duas *Activities* configuradas para receberem comportamento sensível ao contexto.

```

1  {
2    "activities": [
3      {
4        "name": "package.name.awarenesslib.MainActivity",
5        "fences": [
6          {
7            "fenceName": "Headphone",
8            "fenceAction": "action1",
9            "fenceType": "Headphone",
10           "fenceMethod": "Headphone.DURING",
11           "params": {
12             "headphoneState": 1
13           }
14         }
15       ]
16     },
17     {
18       "name": "package.name.awarenesslib.ReconActivity",
19       "fences": [
20         {
21           "fenceName": "Nome Qualquer",
22           "fenceAction": "complex_action",
23           "fenceType": "DetectedActivity",
24           "fenceMethod": "DetectedActivity.STARTING",
25           "params": {
26             "activityTypes": [0,1,2,7,8]
27           }
28         }
29       ]
30     }
31   ]
32 }
33
34 }

```

Figura 2: Exemplo de documento gerado pelo configurador web do EasyContext

O campo *name* é o nome da *activity* juntamente com o pacote do projeto. O campo *fences* é um arranjo de uma ou mais regras de contexto que essa *activity* possa ter. Em cada regra de contexto é necessário informar um nome único (*fenceName*), uma ação para ser executada (*fenceAction*) e a condição para ser observada, representada pelos campos *fenceType*, *fenceMethod* e *params*. Os campos *fenceType* e *fenceMethod* definem uma condição contextual

análoga à encontrada na Google Awareness API, enquanto *params* são parâmetros que configuram essas funções.

3.2 Biblioteca Java

Ao observar a Figura 1, nota-se que a biblioteca Java é a parte central do *framework*, agindo como interpretador de comandos de alto nível gerados pelo Configurador Web e servindo como camada de abstração das funcionalidades do Google Awareness API. Assim como a Awareness API, a biblioteca é dividida em duas partes: Snapshot e Fence. A classe Snapshot é composta de métodos estáticos para aquisição do contexto atual, além de atalhos para alguns métodos da Awareness API. Em termos de padrões de projeto, funciona como um *Façade*, incorporando todos os métodos da Snapshot API.

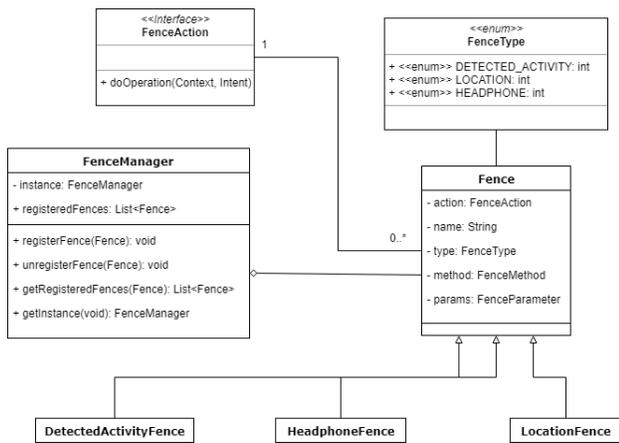


Figura 3: Diagrama de classe simplificado da EasyContext Fence

A parte Fence do EasyContext possui outras classes de apoio, como as mostradas na Figura 3. A interface FenceAction serve como um facilitador para escrita de ações de resposta à mudança de contexto. O desenvolvedor pode implementar esta classe e sobrescrever o método *doOperation* para especificar a ação, de forma a favorecer o reúso de ações. A classe Fence serve como uma representação de uma regra de contexto genérica. Sua instanciação é apenas utilizada pelo configurador de regras contextuais. As classes de apoio para Fence têm como função especificar a regra de contexto em uma forma mais amigável, categorizando e encapsulando os métodos de aquisição de contexto às categorias corretas. Cada classe de apoio possui um Builder para a especificação de parâmetros extras. O FenceManager é um *Singleton* para a gerência de regras contextuais. Esta classe tem como objetivo registrar e remover regras contextuais da pilha do Awareness API, além de realizar o registro dos *Broadcasts* em paralelo, fora do domínio do programador. Devido a este controle, o código que o desenvolvedor deve escrever é reduzido, pois não há mais a necessidade de reescrever diversas linhas de código que cuidam do gerenciamento destes *Broadcasts*.

A classe Configurador configura determinada tela do aplicativo e a deixa preparada para ser sensível ao contexto de forma automática. Contudo, há um leitor JSON que lê o documento gerado pelo configurador web e cataloga todas as regras de contexto em um formato interpretável pelo Google Awareness API. Este documento deve

estar localizado na pasta */res/raw/*. O Configurador também se utiliza de gerência de regras de contexto utilizando o FenceManager, que registra e remove regras quando sua observação não é mais necessária para a sensibilidade ao contexto. Um exemplo em que este fenômeno acontece é quando o usuário troca de atividade dentro da aplicação. Dependendo da *Activity* em que o Configurador é invocado, é catalogado registrado as regras que estão relacionadas com tal *Activity*. Abaixo se encontram duas listagens comparando as duas abordagens, a primeira utilizando o *framework* EasyContext para a configuração de regras de contexto. O documento de regras é o mesmo apresentado na Figura 2. A segunda listagem utiliza a API padrão do Google Awareness.

Listagem 1: Exemplo de código do Configurador EasyContext

```

1 @Override
2 protected void onCreate(Bundle
3     savedInstanceState) {
4     super.onCreate(savedInstanceState);
5     setContentView(R.layout.activity_main);
6     Map<String, FenceAction> actions = new
7         HashMap<>();
8     actions.put("action1", new MyCustomAction());
9     Configurador.init(MainActivity.this, actions);
10 }
    
```

Listagem 2: Exemplo de código padrão sem EasyContext

```

1 @Override
2 protected void onCreate(Bundle
3     savedInstanceState) {
4     super.onCreate(savedInstanceState);
5     setContentView(R.layout.activity_main);
6     BroadcastReceiver action = new
7         BroadcastReceiver() {
8         @Override
9         public void onReceive(Context context,
10             Intent intent) {
11             FenceState state =
12                 FenceState.extract(intent);
13             switch (state.getCurrentState()) {
14                 case FenceState.True:
15                     System.out.println("Headphone
16                         connected");
17                     break;
18                 case FenceState.False:
19                     System.out.println("Headphone
20                         disconnected");
21                     break;
22             }
23             registerReceiver(action, new
24                 IntentFilter("headphone"));
25             PendingIntent pi =
26                 PendingIntent.getBroadcast(this, 0,
27                     new Intent("headphone"),
28                     PendingIntent.FLAG_CANCEL_CURRENT);
29             AwarenessFence fence =
30                 HeadphoneFence.during(
31                     HeadphoneState.PLUGGED_IN
32                 );
33         }
34     };
35 }
    
```

```

22     Awareness . getFenceClient ( this )
23     . updateFences ( new
24         FenceUpdateRequest . Builder ()
25         . addFence ( "headphone" , fence , pi ) . build () ) ;
26     }
27 }

```

Ainda que o programador dedique uma classe que implemente `BroadcastReceiver` no seu projeto, ele deve realizar o registro do mesmo no escopo onde deseja sensibilidade ao contexto, prejudicando a legibilidade do código.

4 PROVA DE CONCEITO

Foi feita uma pequena aplicação como prova de conceito para ilustrar o uso do `EasyContext` (Figura 4). A aplicação utiliza o `EasyContext` para que o telefone se comporte nos seguintes cenários: vibrar quando o usuário estiver andando ou conectar um fone de ouvido e ligar a lanterna quando o usuário estiver andando e com o fone de ouvido conectado ao mesmo tempo. A Listagem 3 exibe um trecho de seu código.

Listagem 3: Prova de Conceito do `EasyContext`

```

1  protected void onCreate ( Bundle
2      savedInstanceState ) {
3      super . onCreate ( savedInstanceState ) ;
4      setContentView ( R . layout . activity_main ) ;
5      HeadphoneFence pluggedIn = new
6          HeadphoneFence ( "plugged" ,
7          HeadphoneMethod . HEADPHONE_PLUGGING_IN ,
8          new VibeAction () ,
9          null ) ;
10     DetectedActivityFence walking = new
11         DetectedActivityFence ( "walking" ,
12         DAMethod . DA_STARTING , new VibeAction () ,
13         new DetectedActivityParameter
14             . Builder ()
15             . addActivityType
16             ( DetectedActivity . WALKING )
17             . build () ) ;
18     AndFence and = new AndFence ( "composition" , new
19         LightAction ( this ) , pluggedIn , walking ) ;
20     FenceManager manager =
21         FenceManager . getInstance ( this ) ;
22     manager . registerFence ( pluggedIn ) ;
23     manager . registerFence ( walking ) ;
24     manager . registerFence ( and ) ;
25 }

```

5 CONSIDERAÇÕES FINAIS

Este artigo propõe uma abordagem para facilitar o desenvolvimento de aplicativos CAM usando a `Awareness API`, o `framework EasyContext`, que encapsula a complexidade da configuração e uso da API. A abordagem ainda é passível de melhorias. A primeira é a falta de suporte para criação de regras de contexto agregadas (composições de Fences) utilizando o configurador web. A forma de se representar estas regras podem ser muito complexas e esta complexidade pode influenciar no tamanho do arquivo de configuração. O `Configurator`

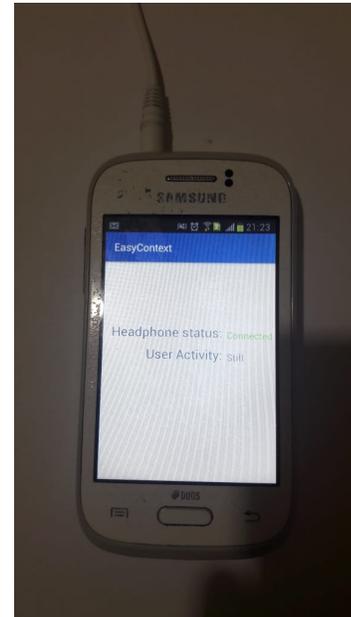


Figura 4: Prova de Conceito em um hardware Android

Java ainda leva um tempo para fazer a leitura e registro das regras de contexto. Uma solução seria otimizar partes do documento para serem lidas apenas pelas `Activities` as quais estas regras se referem. Uma avaliação baseada em técnicas de Engenharia de Software Experimental também está em andamento para comprovar os ganhos da abordagem.

REFERÊNCIAS

- [1] Paolo Bellavista, Antonio Corradi, Mario Fanelli, and Luca Foschini. 2012. A Survey of Context Data Distribution for Mobile Ubiquitous Systems. *ACM Comput. Surv.* 44, 4, Article 24 (Sept. 2012), 45 pages.
- [2] Anind K. Dey. 2001. Understanding and Using Context. *Personal Ubiquitous Comput.* 5, 1 (Jan. 2001), 4–7. <https://doi.org/10.1007/s007790170019>
- [3] Paulo A. S. Duarte, Felipe M. Barreto, Francisco A. A. Gomes, Windson Viana, and Fernando A. M. Trinta. 2015. CRITICAL: A Configuration Tool for Context Aware and mobile Applications. In *2015 IEEE 39th Annual Computer Software and Applications Conference*, Vol. 2. 159–168.
- [4] Eric Freeman, Elisabeth Robson, Bert Bates, and Kathy Sierra. 2004. *Head First Design Patterns: A Brain-Friendly Guide*. O'Reilly Media, Inc.
- [5] José R. Hoyos, Jesús García-Molina, and Juan A. Botia. 2013. A Domain-specific Language for Context Modeling in Context-aware Systems. *J. Syst. Softw.* 86, 11 (Nov. 2013), 2890–2905.
- [6] Fei Li, S. Sehic, and S. Dustdar. 2010. COPAL: An adaptive approach to context provisioning. In *2010 IEEE 6th International Conference on Wireless and Mobile Computing, Networking and Communications*. 286–293.
- [7] Marcio E. F. Maia, Andre Fonteles, Benedito Neto, Romulo Gadelha, Windson Viana, and Rossana M. C. Andrade. 2013. LoCCAM - Loosely Coupled Context Acquisition Middleware. In *Proceedings of the 28th Annual ACM Symposium on Applied Computing (SAC '13)*. ACM, New York, NY, USA, 534–541.
- [8] André C. Santos, João M. P. Cardoso, Pedro C. Diniz, and Diogo R. Ferreira. 2013. Specifying Adaptations through a DSL with an Application to Mobile Robot Navigation. In *2nd Symposium on Languages, Applications and Technologies, SLATE 2013, June 20-21, 2013 - Porto, Portugal*. 219–234.
- [9] O. B. Sezer, E. Dogdu, and A. M. Ozbayoglu. 2018. Context-Aware Computing, Learning, and Big Data in Internet of Things: A Survey. *IEEE Internet of Things Journal* 5, 1 (Feb 2018), 1–27. <https://doi.org/10.1109/JIOT.2017.2773600>
- [10] Ö. Yürür, C. H. Liu, Z. Sheng, V. C. M. Leung, W. Moreno, and K. K. Leung. 2016. Context-Awareness for Mobile Sensing: A Survey and Future Directions. *IEEE Communications Surveys Tutorials* 18, 1 (Firstquarter 2016), 68–93. <https://doi.org/10.1109/COMST.2014.2381246>

Towards Data Transmission Through Inaudible Sound in Ginga-NCL

João Victor G. de S. Nunes, Álan L. V. Guedes, Guilherme F. Lima, Sérgio Colcher

Pontifical Catholic University of Rio de Janeiro

Rio de Janeiro, Rio de Janeiro

jvgirard@telemidia.puc-rio.br

ABSTRACT

In this paper, we report our efforts to add support for data transmission through inaudible sound to the Ginga-NCL Digital TV middleware. We present an algorithm for encoding a bitstream in an inaudible audio signal, and to do so reliably on consumer-grade hardware. We also discuss two attempts to implement this algorithm in NCL, the language in which Ginga-NCL applications are written. The first attempt was to transmit prerecorded inaudible audio signals in a Ginga-NCL-compatible set-top-box. And the second attempt was to use NCLua to generate at runtime the inaudible audio signal. For the second attempt we extended NCL with a novel media object type, called *SigGen*, which can be used to generate arbitrary audio signals. In the paper, we describe in detail the implementation of *SigGen* and the result of these attempts.

KEYWORDS

Inaudible sound; Ginga-NCL; Nested Context Language; NCL

1 INTRODUCTION

Current consumer devices, such as TVs, smartphones, smartwatches, etc., often come with built-in support for radio-frequency (RF) wireless communication. This support requires RF hardware and depends on technologies such as Bluetooth, IEEE 802 WiFi, or 4G mobile networks. There are situations, however, where the RF data network is not available, or where it is more advantageous to use some other form of wireless communication. One such alternative is communication through inaudible sound.

The same consumer devices that come with RF support usually come equipped with speakers and microphones which enables them to produce and detect inaudible sound. By inaudible sound we mean sound waves in frequencies that cannot be heard by humans—usually those above 19–22 kHz. There are many works in the literature that show that communication through inaudible sound is feasible on consumer-grade PCs, laptops, and smartphones [3, 4, 6]. Here we are mainly concerned with the applicability of this technique to digital TV scenarios, in particular, to those scenarios where the TV broadcasts data to nearby devices through inaudible sound.

In this paper, we report our ongoing efforts to enable inaudible sound applications in the Ginga-NCL Digital TV middleware [1]. These efforts, so far, consisted of two steps. First, we came up with a method for encoding a bitstream in an inaudible audio signal, and

to do so in a way that the resulting signal could be generated and decoded reliably on consumer-grade hardware.

After that, in a second step, we tried to implement this encoding algorithm in NCL, which is the language in which Ginga-NCL applications are written. NCL is a declarative language with no support for audio synthesis. So, our first attempt was to record the signals generated by our algorithm into raw audio files; then we combined these files in an NCL document and tried to play the document in a Ginga-NCL-compliant set-top box. Unfortunately, due to hardware limitations, that didn't work.

In face of this, we decided to extend the NCL language with a new type of media object, called *SigGen*, which can generate arbitrary audio signals. We implemented the *SigGen* object and its underlying player in PUC-Rio's NCL player, using the GStreamer multimedia framework. Then we implemented our encoding algorithm in an NCLua script (the scripting language of NCL) which was used to drive *SigGen* objects. Unfortunately, that also didn't work. (In PUC-Rio's implementation, the rate at which changes occur within an NCL document is limited by the display frame-rate—about 60Hz—which is too low for the correct operation of our encoding algorithm.)

The rest of the paper is organized as follows. In Sections 2 and 4, we describe our main contributions: the encoding algorithm and the *SigGen* media object. In Sections 3 and 4, we give a detailed account of our attempts to implement the proposed algorithm in NCL and of the lessons we learned in the process. In Section 5, we compare our encoding algorithm with similar algorithms we found in the literature. Finally, in Section 6, we present our conclusions and future work.

2 THE ENCODING ALGORITHM

We will describe our encoding algorithm in two parts. In the first part, Section 2.1, we discuss how we selected inaudible frequencies that combined minimize the audible background noise. In the second part, Section 2.2, we discuss the encoding algorithm together with its transmission and reception parts.

The hardware we used in the development of this algorithm is illustrated in Figure 1a. It consisted of a PC with speakers and a smartphone. The PC had an Intel Core i5 (6th gen.) processor, 16 GB of RAM memory, Edifier XM2PF speakers, and run Ubuntu 16.04 LTS. The smartphone was a Samsung Galaxy S8 running Android 8.0 Oreo.

We implemented the transmission part of the algorithm in the GStreamer multimedia framework. The GStreamer pipeline we used for the tests consisted of three elements connected in series: *AudioTestSrc* to generate frequencies; *AudioConvert* to convert raw audio buffers to samples; and *AutoAudioSink* to send the raw samples

In: XV Workshop de Trabalhos de Iniciação Científica (WTIC 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pósteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.

© 2018 SBC – Sociedade Brasileira de Computação.

ISBN 978-85-7669-435-9.

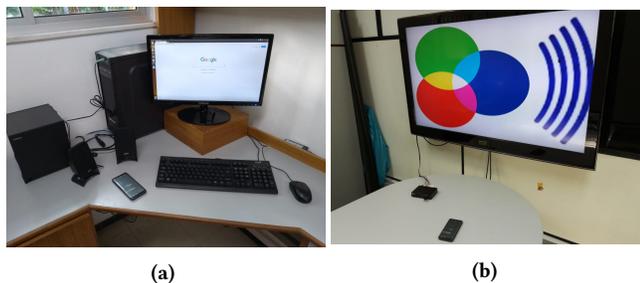


Figure 1: (a) PC environment and (b) TV environment (set-top box); both broadcasting to an Android smartphone.

to the audio driver. For reception, we used the Android AudioRecord API¹ to analyze the audio spectrum by taking samples.

2.1 Frequency Selection

To explain our selection process, we first need to introduce the concepts of *guard band* and *beat frequencies*. The guard band is the gap in the audio spectrum between two data-carrying frequencies. This gap is introduced to ensure the unambiguous reception of the data carried by contiguous frequencies. Sometimes, however, the combination of data-carrying frequencies and guard band may generate beat frequencies. These beats happen, for instance, when two frequencies interfere with each other. Such interference can generate a new frequency with a periodic variation of volume and a rate that is equal to the difference of the two initial frequencies. Beat frequencies are undesirable because they can generate noise in frequencies which can be heard by humans.

We chose frequencies based on experiments we did in a quiet room. The goal was to select the maximum number of frequencies which generated fewer *beats* and had a good *guard band*. Figure 2 depicts the spectrum visualization of four experiments using the Spectrum Analyzer². In the figure, the cyan rectangle highlights beats.

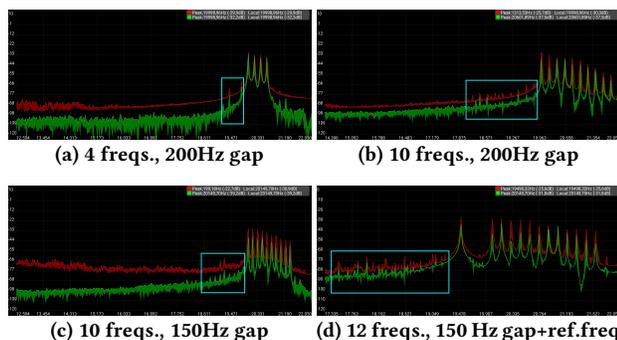


Figure 2: Spectrum analysis of four experiments with frequencies ranging from 17,335Hz to 22,050Hz.

In our definition of “inaudible”, we chose the starting point of frequencies to be at least above 19000Hz, since some people can hear

¹developer.android.com/reference/android/media/AudioRecord

²play.google.com/store/apps/details?id=com.raspw.SpectrumAnalyzer

sounds in frequencies below this value. Fewer can hear above it, so it is viable to work in this area of the spectrum. The reception hardware defines our maximum frequency to be around 22000Hz, which seems to be the operational limit for consumer-grade hardware.

In the first experiment (Figure 2a), we used 4 frequencies with a 200Hz guard band. These four frequencies generated fewer beats but also carried fewer data (assuming one bit per frequency). In the second experiment, we used 10 frequencies (Figure 2b) with the same 200Hz guard. Although more data was transmitted, these 10 frequencies increased the width and amplitude of the beats. In the third experiment we changed the guard band to 150Hz (Figure 2c) and kept the same 10 frequencies. This retained a good reception and reduced both problems caused by the added frequencies with a larger guard band.

We found our best condition using 13 frequencies (Figure 2d). In this experiment, the first frequency starts at 19500Hz and is separated from the next frequency by a gap of 500Hz. This first frequency is intended to be easily captured as a reference frequency. The greater initial gap between the reference frequency and the others reduces the probability of interference between them. The remaining frequencies are located from 20000Hz to 21650Hz and are separated by a 150Hz guard band. The inclusion of the reference frequency added some beats with very low amplitudes in the high-frequency area, making it possible but hard to be heard by users, but it improved reception without affecting the bit-rate.

2.2 The Algorithm

In our algorithm, the reference frequency is at 19500Hz; this frequency signals the presence of the transmission but carries no data itself. The remaining frequencies carry one bit each, which are to be interpreted in the little-endian format. A message is transmitted by varying the amplitude of these data-carrying frequencies over time, i.e., by toggling them up and down. If the frequency being analyzed is below -60dB the message as a whole is considered invalid, provided that a single misread bit could interfere on the message as a whole.

Thus, above -60dB we set a new amplitude threshold. An amplitude above this new threshold corresponds to a bit 1 and, otherwise, to a bit 0. This threshold is determined dynamically accordingly to the current amplitude of the reference frequency. This is done because the distance between the speaker and the smartphone affects the amplitude of the frequencies when is received by the microphone. The fixed threshold for invalid bits (-60 dB) should not be lower, due to other ambient sounds possibly be mistakenly considered valid or higher so that the dynamic threshold has a wide enough area for the bit 0 amplitude reception.

The reception is done in three steps: sample capture, conversion and analysis (we have not implemented a synchronization method). First we capture 44100 samples per second, as it is recommended to function in all Android devices in the API’s documentation and it suits our needs. By the Nyquist Theorem, this provides us, at most, coverage for 22050Hz frequency. Each sample takes around 0.5ms to be captured. Then we use a FFT (Fast Fourier Transform) algorithm with a buffer size of 512 to convert the raw sound sample to values in dB. Finally, we analyze the amplitude of each frequency comparing

itself to the reference frequency amplitude. One frequency is a bit 1 if it has the same or greater amplitude than the reference frequency.

After all the samples of the bits of the message are collected during its 100 ms period, the previous amount of 1's and 0's of that particular message are analyzed. It is determined which one has at least 2/3 of the total count of samples. In the end, it is set if that bit is a 1 or 0 and then compose the message's eventual outcome.

3 THE SET-TOP BOX EXPERIMENT

Before evaluating our approach in the TV environment (Figure 1b), we experimented on the PC environment (Figure 1a). In particular, we checked if the same signal which was generated in real-time by our GStreamer pipeline could be transmitted by pre-rendered audio samples—and the result was positive. The pre-rendered message was received just as the real-time one. On average, in the PC environment we achieved the same bit-rate using pre-rendered samples. We did this experiment because although TV environment has no support for audio synthesis it can play audio files. So, one approach to inject an inaudible audio signal in the TV environment is through an application which uses such pre-rendered samples.

The Brazilian Digital TV system uses the Ginga-NCL middleware for application development. Those applications are written in the NCL language [1]. Figure 3 depicts an initial version of this application. Each frequency uses a `<media>` element for each pre-rendered sample. The volume property defines the bit value.

```
<media id="freq_ref" src="19500.mp3">
  <property name="volume" value="100%">
</media>
<media id="freq_1" src="20000.mp3">
  <property name="volume" value="30%">
</media>
```

Figure 3: Part of the NCL code using pre-rendered samples.

The TV environment we used (depicted on Figure 1b) consisted of a TV plus a TS 2017 set-top box. This set-top box was distributed by the Brazilian government during the switch-off of the analog TV signal. The deployment of our application in this environment led us to discover a *decoding bottleneck*. Although the TS 2017 set-top box is compliant with the SBTVD [1] requirements, it has only one decoder for video and one decoder for audio, which means that it cannot play more than one audio file at a time. This limitation prevented us from using the pre-rendered audio samples.

One possible workaround to this bottleneck is to encode the signal into the audio transmitted by the broadcaster. The audio signal can then be mixed with the video being broadcast and received in all television sets, including the sets of those users not interested on it. But regarded that only high frequency sounds are being generated to transmit the message, those users do not need to worry about it.

4 THE SIGGEN MEDIA OBJECT

With the failure of the set-top box experiment, our next attempt was to extend Ginga-NCL with support for audio synthesis. We did this by integrating our GStreamer pipeline into PUC-Rio's NCL player [7], as a new media player, called SigGen. More precisely, we extended NCL with new type of `<media>` element, called

`x-ginga-siggen`. Figure 4 presents a NCL application that creates such `<media>` elements. The `freq` and `volume` properties of these elements are used to synthesize a pure frequency.

```
<media id="freq_ref" type="application/x-ginga-siggen">
  <property name="freq" value="19500"/>
  <property name="volume" value="1"/>
</media>
<media id="freq_1" type="application/x-ginga-siggen">
  <property name="freq" value="20000"/>
  <property name="volume" value="0.3"/>
</media>
```

Figure 4: SigGen media object declarations in NCL.

We extended our initial pipeline to also present a visual representation of each frequency. This way a visual representation of the frequencies can be shown on screen. The complete GStreamer pipeline we used is depicted in Figure 5. Note that the SigGen player does not require an audio decoder: it generates raw audio data which can be passed directly to the audio driver.

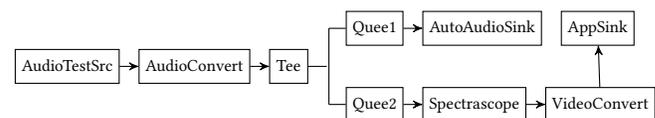


Figure 5: Player SigGen pipeline.

With the SigGen player working, we moved on to implement our encoding algorithm using an NCLua script. The idea was to use such script to change the properties of SigGen media elements (one per frequency) dynamically via the NCLua event .timer API. However, we run into another problem. In PUC-Rio's NCL player, the speed at which the event-loop runs is limited by the monitor frame-rate (usually, 60Hz). This upper limit is too low for the correct implementation of our encoding algorithm, which requires a precision of the order of microseconds.

One way to work around this problem is what is called, the *use case Timeline*. This use case is not intended to have a high bit-rate requirement to transmit messages. Instead it focus on transmitting time-stamps with precision greater than Ginga's event loop, so the bottleneck is surpassed by making the algorithm slower. To send time-stamps during a TV show, for instance, we may use the first four frequencies to determine the show and the remaining 8 frequencies to carry the time-stamps since the show began. For example, a pre-downloaded image or video propaganda can be presented at some specific moment of the TV show on the smartphone. Two videos were made to present a simulation of this propaganda scenario. On video³ the smartphone screen is shown and on video⁴ the transmission as a whole.

5 RELATED WORK

Other works also share our motivation in terms of data transmission using inaudible sound. Jang et al. [5] and Bang et al. [2] focus on sending only one message during all transmission, whereas

³imgur.com/a/FUov3mu

⁴imgur.com/a/MirE3sy

Table 1: Comparison summary.

Feature	Jang [5]	Bang [2]	Lakhwani [8]	Tsugawa [9]	This work
Number of data/total frequencies	32/40	32/40	2/2	8/20	12/13
Sampling rate (Hz)	48000	44100	44100	44100	44100
FFT sample size	32768	8192	1024	512	512
Frequencies used (kHz)	18 to 24	18 to 22	20.7 and 21	17.959 to 20.973	19.5 and 20 to 21.6
Dynamic bit threshold	yes	no	yes	yes	yes
Error handling method	CRC 8 bits	CRC 8 bits	no	Hamming Code (7,4)	Invalid bit
Send time (ms)	92	500	105	150	100+300 (data+silence)
Bit-Rate (bps)	347 (same msg)	64 (same msg)	19	53	30
Success rate	99.8%	97.5%	76%	93%	60%
Synchronization method	no	no	handshake	preamble detection	no
Noise when redone with GStreamer	yes	yes	no	yes	no

Lakhwani et al. [8] and Tsugawa et al. [9] focus on sending a sequence of messages. We discuss each of these works next. Table 1 presents a summary of their main features.

Jang et al. [5] use 40 evenly divided frequencies in a flexible range between 18000Hz and 24000Hz. They use 8 bits for CRC (cyclic redundancy correction) and transmit each piece of data for 92ms, achieving a bit-rate of 347bps. Their main limitation is the fact that they use 48kHz as sampling rate. This sampling rate increases the number of frequencies that can be used, but is not recommended for reception on Android smartphones.

Bang et al. [2], similarly to Jang et al., use 40 frequencies between 18050Hz and 21950Hz with a guard band of 100Hz. They also use 8 bits for CRC and transmit each piece of data for 500ms, achieving bit-rate of 64bps. Their main limitation is that they do not have a dynamic bit threshold to handle distance variation between the transmitter and the receptor.

Lakhwani et al. [8] use only two frequencies, 21000Hz and 20700Hz, the former representing a bit 1 and the latter representing a bit 0. They transmit each piece of data for 105ms achieving a bit-rate of 19bps. To handle the synchronization of messages, their transmitter also have a microphone and performs a handshake with the receiver. Lakhwani et al. are not concerned with improving the bit-rate or correcting errors.

Tsugawa et al. [9] use 20 frequencies between 17959Hz and 20973Hz. Six of these are used for Hamming Code error correction, two for dynamic bit threshold, and two for synchronization. The synchronization bits are used before the data bit analysis happens at the start of each transmission. Their algorithm waits for two successful transmissions of the synchronization bits in order to guarantee the correct start of the data bits analysis. Tsugawa et al. transmit each piece of data for 150ms achieving a bit-rate of 53bps. However, they work only with pre-rendered audio samples generated on MATLAB.

Our transmission algorithm achieves a bit-rate of 30bps. On the PC environment described earlier, it has achieved an average rate of successful reception for the whole message of 60% and for the bits individually of 76%. The algorithm with the highest transmission rate in the works we are considering is the one from Tsugawa et al. This algorithm also has a good success rate. We used our GStreamer pipeline to test the frequency selections of all algorithms discussed in this section. Except for Lakhwani et al., which only uses two frequencies, all of the selections generated frequency beats and noise.

Comparing the number of frequencies we used (13) to the number used in other works, e.g., 20 in Tsugawa et al., our work indeed uses less frequencies, but retains no audible noise reproduction coming from undesired beat frequencies while using the GStreamer pipeline with our frequencies placement. An increase in the number of frequencies to 20 could be attempted in two ways. One way is to widen the area of the spectrum used by the reallocation of the reference frequency, but this could increase the amount of noise generated. Another way is to decrease the gap between the frequencies and increase the FFT sample size. This, however, would slow down the capture speed of the Android devices which would reduce the bit-rate.

6 FINAL REMARKS

This paper presented lessons learned from our experiments to transmit data through inaudible sound in the PC and TV environments. Although related work achieves higher bit-rates, our contribution is in the discussion of the problems and workarounds for making the technique feasible on the digital TV environment.

As future work, we intent to improve the use of our data-carrying frequencies in order to increase the bit-rate achieved by our algorithm. We also intend to extend the algorithm to handle errors and to improve the transmission-reception synchronization.

REFERENCES

- [1] ABNT 15606-2. 2007. Digital Terrestrial TV – Data Coding and Transmission Specification for Digital Broadcasting – Part 2: Ginga-NCL for Fixed and Mobile Receivers: XML Application Language for Application Coding. (2007).
- [2] Green Bang, Myoungbeom Chung, and Ilju Ko. 2016. Data communication method based on inaudible sound at near field. (2016), 4.
- [3] Luke Deshotels. 2014. Inaudible Sound As a Covert Channel in Mobile Devices. In *Proc. 8th USENIX Conf. Offensive Technologies*.
- [4] Michael Hanspach and Micahel Goetz. 2013. On Covert Acoustical Mesh Networks in Air. *J. Communications* 8, 11 (2013). DOI: <http://dx.doi.org/10.12720/jcm.8.11.758-767>
- [5] Insu Jang, Myoungbeom Chung, and Hyunseung Choo. N/A. Reliable Short-Distance Data-Transmission Mechanism Using Inaudible High-Frequency Sound. (N/A), 10. https://insujang.github.io/assets/pdf/research_paper_data_communication.pdf
- [6] Soonwon Ka, Tae Hyun Kim, Jae Yeol Ha, Sun Hong Lim, Su Cheol Shin, Jun Won Choi, Chulyoung Kwak, and Sunghyun Choi. 2016. Near-ultrasound Communication for TV’s 2nd Screen Services. In *Proc. ACM 22nd Ann. Int. Conf. on Mobile Computing and Networking*.
- [7] TeleMidia Lab. 2018. (2018). <https://github.com/TeleMidia/ginga>
- [8] Sahil Lakhwani, Nishant Pardamwar, and Nikhil Khewalkar. 2015. High Frequency Sound Based Device Communication. *IJARCCCE* 4, 3 (March 2015).
- [9] Hiroaki Tsugawa and Masakatsu Ogawa. 2017. Proposal of Ultrasonic Communication Method and Its Application to Position Estimation System. *J. Signal Processing* 21, 4 (2017).

MathQuiz

A Game App for M-Learning

José Marcelo Silva Santiago
Universidade Estadual do Ceará -
UECE
Limoeiro do Norte, Ceará
marcelo.gmcf@gmail.com

Monck Charles Nunes de
Albuquerque
Universidade Estadual do Ceará -
UECE
Fortaleza, Ceará
monck.albuquerque@uece.br

Fernanda Kécia de Almeida
Universidade Estadual do Ceará -
UECE
Limoeiro do Norte, Ceará
fernanda.kecia@uece.br

Francisco Ranulfo Freitas
Martins
Universidade Estadual do Ceará -
UECE
Limoeiro do Norte, Ceará
ranulfo.freitas@uece.br

Yuri Soares de Oliveira
Instituto Federal de Educação, Ciência
e Tecnologia do Ceará - IFCE
Fortaleza, Ceará
yuri.oliveira@ppgcc.ifce.edu.br

ABSTRACT

This paper presents the MathQuiz application as an option to gamify some mathematics classes in elementary education II. Basically, MathQuiz is a question and answer quiz where the teacher can register questions. MathQuiz was developed for android and can be easily installed on mobile phones. MathQuiz was created in a TCC job, has been tested and perfected, and is ready to be used in classrooms.

KEYWORDS

Game, M-learning, Math

1 INTRODUÇÃO

O desenvolvimento de *softwares* alterou significativamente o modo e estilo de vida das pessoas ao redor do mundo. Desde o seu surgimento, o software se estabeleceu como a mais importante tecnologia no cenário mundial, algo inimaginável há 60 anos [16]. Nas últimas décadas, a capacidade das tecnologias Digitais da comunicação e informação (TDIC)¹ aumentou de forma exponencial e vem revolucionando vários setores, inclusive a educação [1].

A disseminação dos *softwares* e a introdução a cultura digital fez com que a maioria das pessoas tivessem seus hábitos alterados [8]. Este panorama ocorre principalmente quando falamos dos nascidos digitais, uma geração de crianças e adolescentes que já nasceram num contexto de imersão ao ambiente online. Novas gerações são adeptas das funcionalidades da mídia digital, permitindo novas ações comunicativas descentradas e *mobiles* [5].

¹Segundo [11], TDIC são tecnologias que têm o computador e a Internet como instrumentos principais e se diferenciam das Tecnologias de Informação e Comunicação (TIC) pela presença do digital.

In: XV Workshop de Trabalhos de Iniciação Científica (WTIC 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pósteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.

© 2018 SBC – Sociedade Brasileira de Computação.
ISBN 978-85-7669-435-9.

Atualmente, os dispositivos móveis, como *smartphones* e *tablets*, se tornaram itens quase indispensáveis no cotidiano das pessoas. Nesses dispositivos, um dos sistemas operacionais mais utilizados é o *Android* da Google[6]. Hoje na sala de aula quase todos os alunos têm um celular que pode ser utilizado para auxiliar no processo de aprendizado do aluno, portanto a educação presencial pode ser significativamente modificada com base na tecnologia, pois permite que as escolas e universidades saiam do limite geográfico para trocar informações, facilitando as pesquisas e construindo o conhecimento.

A educação, como o processo de aprender, pode ser otimizada pela possibilidade de integração de várias mídias, acessando-as tanto em tempo real como assincronicamente, assim como também pela facilidade de pôr em contato educadores e educandos [12].

Assim, como várias outras instituições e meios de comunicação estão se adequando para atender às necessidades das TDICs, a escola, enquanto espaço propício para integrações, também precisa entrar na chamada cultura da convergência, trazendo para suas práticas os benefícios da tecnologia [10].

Partindo disso, propusemos uma análise desses benefícios das TDICs para ajudar no processo de aprendizagem, olhando especificamente para as dificuldades enfrentadas na aprendizagem de Matemática nas séries do Ensino Fundamental II. A proposta foi desenvolver um jogo para celular, que o professor possa usar na sala de aula, permitindo que os alunos superem suas dificuldades e possam compreender melhor os assuntos sobre Matemática.

Na literatura nacional são encontrados vários trabalhos que relatam o uso de jogos na educação, porém são softwares destinados a computadores. Quando trazemos essa abordagem para o contexto de dispositivos móveis (*smartphones* e *tablets*), percebemos que ainda é tímido o número de trabalhos para essa área em específico, levando em consideração o crescimento no uso desses dispositivos móveis [15].

A metodologia utilizada nessa pesquisa, parcialmente adaptada de [7] foi focada nos alunos do Ensino Fundamental II para a experimentação de um Jogo Educativo. Foi proposto aos alunos a utilização do Jogo para responder a perguntas de matemática voltadas ao seu

ano de ensino. Foram coletadas, por meio de questionários, algumas informações após o uso do Jogo Educativo.

Para o desenvolvimento do Jogo Educativo foi utilizado o Modelo incremental no processo de desenvolvimento de *softwares*, foram definidos os requisitos funcionais e não funcionais como também a relação das regras de negócios. Foram montados os diagramas de caso de uso, o diagrama de classe e seus atributos para a confecção do banco de dados.

O artigo está organizado em seis seções. A seção 1 apresenta uma introdução da utilização das tecnologias na educação e os objetivos da pesquisa. A seção 2, apresenta jogos educacionais. Na seção 3, alguns trabalhos relacionados com o objeto de pesquisa. Na seção 4, apresenta-se o Jogo Educativo MathQuiz, descrevendo sua dinâmica de funcionamento. Na seção 5, realiza-se uma análise da usabilidade para validação do uso do Jogo. A última seção apresenta as conclusões e sugestões para trabalhos futuros.

2 JOGOS EDUCACIONAIS

No ambiente educacional é importante estimular o acesso a outros meios que possam dar apoio no processo de ensino e aprendizado. Segundo [9], os Jogos Educativos auxiliam na construção do conhecimento e aprendizagem.

Segundo [13], aprendizagem é vista por muitas pessoas como um "dever" para as crianças. Esta visão crítica faz com que a aprendizagem seja encarada como um trabalho. Entretanto, quando uma atividade envolve prazer, diversão, motivação, interesse e paixão, o indivíduo é capaz de dedicar a ela uma grande parte do tempo e esforço. Como jogos digitais envolvem muitos fatores motivacionais, o seu uso poderia ser estimulado por educadores para auxiliar na aprendizagem de conteúdos difíceis de tratar em salas de aula tradicionais [13].

Aprender através de ferramentas interativas não só estimula o aprendizado, como também durante um jogo o usuário se diverte e se sente desafiado para concluir as missões propostas pelo jogo, sendo uma maneira mais eficaz de ensinar [9, 14].

Além da facilidade em aprender através dos jogos, segundo um estudo da Federação de Cientistas Americanos [17], os jogos facilitam a interação social, pois 70% das crianças e adolescentes estudam em conjunto, resolvendo os problemas nos jogos utilizando estratégias conjuntamente [17].

3 TRABALHOS RELACIONADOS

Na internet, podemos encontrar diversos jogos envolvendo conteúdos de matemática, porém, poucos permitem uma personalização por parte do professor.

O jogo "O homem que calculava" proposto por [13] tem o enredo baseado no livro². Sendo um jogo de perguntas e respostas. Foi dividido em seis estágios para se tornar mais atrativo e motivador, com níveis de dificuldade crescente. Foi desenvolvido para a plataforma *Android* e possui o código aberto.

Outro jogo relacionado é o Math4Mobile [4] que permite aos alunos esboçar gráficos, usando sete ícones. Podem ser usados valores para aumentar e diminuir as funções que mudam a constante

²Livro clássico brasileiro traduzido para o inglês e o espanhol, do escritor Malba Tahan, pseudônimo do professor de Matemática Júlio Cesar de Mello e Souza

de inclinação. O Jogo fornece *feedback* imediato sobre o gráfico desenhado.

Uma proposta educativa é apresentada no jogo M-learnMat [3] que é um modelo pedagógico que visa orientar atividades de *m-learning* (*mobile learning*) em matemática. O mesmo é embasado pela teoria da atividade e tem como foco o ensino superior. Também adota o modelo pergunta e resposta.

[18] realizou um mapeamento esquemática sobre Jogos móveis para o aprendizado de matemática e encontrou 130 publicações. Após a aplicação de um segundo filtro, selecionou 10 artigos que melhor representa o processo de ensino e aprendizagem para matemática. Desse total apenas 25% era destinado ao ensino fundamental e 12,5% era na plataforma *Android*.

O diferencial entre os trabalhos apresentados nesta seção e o Jogo MathQuiz é que ele não é finito, pois este, possibilita ao professor o cadastramento de questões e imagens de acordo com o conteúdo abordado em sala de aula.

4 SOBRE O JOGO MATHQUIZ

Esta seção apresenta o Jogo Educativo MathQuiz e suas principais funcionalidades. A tela da (Figura 1a) é a principal após o carregamento inicial. Nela, o usuário tem a opção de escolher a série que está cursando, constituída por alternativas entre o 6° ao 9° ano do Ensino Fundamental II.

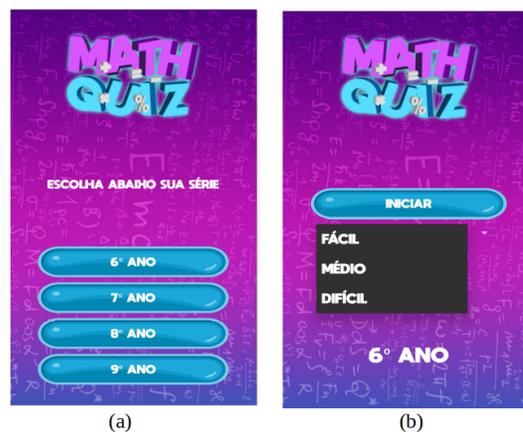


Figura 1: Telas do jogo MatQuiz. (a) Tela inicial com escolha da série (b) tela com a escolha do nível de dificuldade.

O usuário então, será direcionado para uma outra tela contendo a descrição dos níveis das perguntas presentes no Jogo Educativo (Figura 1b). O Jogo conta com três níveis de dificuldades: Fácil, Médio e Difícil.

É apresentada na (Figura 2a) a interface gráfica que o usuário vê e interage. São exibidos: o número da questão, a quantidade de pontos acertados, o cronômetro de resposta, a pergunta exibida no quadro, as opções de respostas, o botão de confirmar a opção selecionada, e o ícone azul de ajuda no canto inferior direito da tela.

Na (Figura 2b) é apresentada a tela em que o usuário seleciona umas das alternativas para a pergunta e o resultado encontra-se errado. Logo, é mostrada na tela uma caixa de diálogo com a mensagem "ERROU!" na cor vermelha, demonstrando que a escolha da



Figura 2: Telas do Jogo Educativo MatQuiz. (a) tela de exibição da questão e as opções de escolha (b) Tela sinalizando a resposta errada.

resposta não foi a correta. Em seguida, a resposta correta é mostrada na (Figura 3a).



Figura 3: Telas do Jogo MatQuiz. (a) tela sinalizando o acerto da questão (b) tela de pontuação.

Por fim, a (Figura 3b) mostra o resultado final após a realização das perguntas do quiz, em que as estrelas representam a conquista em cada nível respondido com 100% de acertos nas perguntas. Abaixo, podemos ver a quantidade de acertos do usuário em cada nível, o tempo total gasto para responder no nível de dificuldade selecionado pelo usuário e a quantidade de pontos totais. No canto inferior esquerdo da tela, é exibido um ícone vermelho com um “x” que tem como função sair da aplicação. Já no canto inferior direito da tela há a opção de reiniciar o quiz.

5 RESULTADOS

O presente estudo investigou a utilização do MathQuiz na Disciplina de Matemática em duas turmas do Ensino Fundamental II. No total de 45 alunos, sendo 21 pertencentes ao 7º ano e 24 do 8º ano.

5.1 Análise da Usabilidade

Foi utilizado o Questionário de Usabilidade (SUS) para a avaliação do Jogo educativo. Para facilitar a compreensão, o grau de usabilidade está indicado no formato numérica, sendo 1 para Discordo completamente, 3 para Neutro, e 5 para Concordo completamente. A seguir, as respostas para algumas das questões contidas no questionário.

Para a frequência do uso foram obtidos os valores de 71,11% dos alunos gostariam de usar o Jogo com frequência, enquanto 11,11% não o utilizariam frequentemente (Figura 4).

Eu acho que gostaria de usar esse sistema com frequência

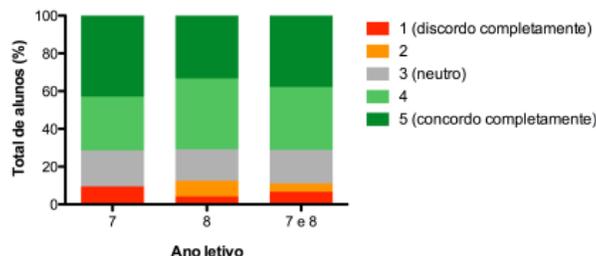


Figura 4: Frequência de Uso do Jogo.

Na (Figura 5), é possível notar que os usuários confirmam que o Jogo é realmente fácil de usar, pois 88,89% deles acham que outras pessoas aprenderiam rapidamente a utilizar o software.

Eu imagino que as pessoas aprenderão como usar esse sistema rapidamente.

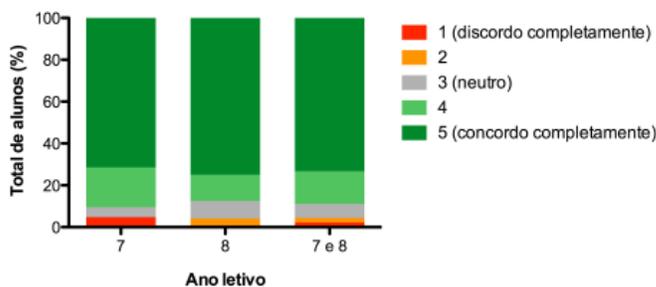


Figura 5: Facilidade de Uso do Jogo.

Para as perguntas sobre complexidade da utilização foi observado que os alunos não acham o sistema complexo para ser utilizado, pois somente 22,22% dos estudantes apontaram que o seu uso demanda uma certa complexidade. Foi percebido que o Jogo Educativo pode ser utilizado de forma autônoma pelos estudantes. A ausência da necessidade de conhecimentos técnicos prévios foi indicado por 60% dos alunos e que os usuários confirmam que o Jogo é realmente fácil de usar, pois 88,89% deles acham que outras pessoas aprenderiam rapidamente a utilizar o software.

6 CONCLUSÃO

Neste artigo foi apresentado o MathQuiz que surgiu como um trabalho de TCC da graduação em computação com o objetivo de ser um Jogo Educativo para o apoio no ensino da matemática para alunos do ensino fundamental.

A validação foi realizada com professores e alunos. O Jogo pode ser utilizado em celular ou *Tablet* e pode ser obtido gratuitamente.

Como trabalhos futuros para o MathQuiz foi proposto:

- Implementar modelo de competições em grupo;
- Desenvolver um Ranking da qualidade das respostas baseado em [2];
- Implementar um Cálculo de eficiência em respostas utilizando Lógica Fuzzy³;
- Realizar um estudo com uma amostra maior e heterogênea de alunos e professores.

Os resultados da pesquisa mostram indícios de que o Jogo Educativo MathQuiz é adequado para o público alvo e que é útil para apoiar o processo de ensino e aprendizagem na disciplina de matemática.

REFERÊNCIAS

- [1] Monck Albuquerque, Gilvandenys Leite Sales, Pedro Rebouças Filho, and Claudio Medeiros. 2017. Avaliação da presencialidade em um fórum LV utilizando Lógica Fuzzy. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, Vol. 28. 1357. <http://dx.doi.org/10.5753/cbie.sbie.2017.1357>
- [2] Leandro Amancio and Carina Friedrich Dorneles. 2017. Towards Recency Ranking in Community Question Answering: A Case Study of Stack Overflow. In *Proceedings of the 23rd Brazilian Symposium on Multimedia and the Web*. ACM, 173–180.
- [3] Sílvia Cristina F Batista, Patricia Alejandra Behar, and Liliana Maria Passerino. 2011. M-learnMat: Aplicação de um Modelo Pedagógico para Atividades de M-learning em Matemática. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, Vol. 1.
- [4] Galit Botzer and Michal Yerushalmy. 2007. Mobile application for mobile learning. In *Proceedings of IADIS International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2007)*. 7–9.
- [5] Belarmino César Guimarães da Costa. 2010. Comunicação e Educação na Era digital: reflexões sobre estética e virtualização. *Revista Comunicação, Mídia e Consumo* 7, 19 (2010).
- [6] Emerson Barbosa da Cunha, Daniella Dias Cavalcante da Silva, and César Rocha Vasconcelos. 2017. Enabling Full Interaction with the Android System and Applications through Speech Recognition. In *Proceedings of the 23rd Brazilian Symposium on Multimedia and the Web*. ACM, 497–500.
- [7] Dayane de Souza, Genarde Trindade, Priscila Silva Fernandes, and Bruno Bonifácio. 2017. alfabetzAR: Uma Aplicação Móvel com base na Realidade Aumentada como Ferramenta de Apoio no Processo de Alfabetização de Portadores de Síndrome de Down. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, Vol. 28. 897.
- [8] Adrián Domínguez, Joseba Saenz-De-Navarrete, Luisde-Marcos, Luis Fernández-Sanz, Carmen Pagés, and José-Javier Martínez-Herráiz. 2013. Gamifying learning experiences: Practical implications and outcomes. *Computers and Education* 63 (2013), 380–392.
- [9] Neusa Nogueira Fialho. 2008. Os jogos pedagógicos como ferramentas de ensino. In *Congresso nacional de educação*, Vol. 6. 12298–12306.
- [10] R Koster. 2004. *A theory of fun*. Paraglyph Press, New York, NY, USA.
- [11] Simão Pedro MARINHO and Wolney Lobato. 2008. Tecnologias digitais na educação: desafios para a pesquisa na pós-graduação em educação. *Colóquio de Pesquisa em Educação* 6 (2008), 1–9.
- [12] José Manuel Moran. 1997. Como utilizar a Internet na educação. *Ciência da informação* 26, 2 (1997).
- [13] José Francisco Barbosa Neto and Fernando de Souza da Fonseca. 2013. Jogos educativos em dispositivos móveis como auxílio ao ensino da matemática. *RENOTE* 11, 1 (2013).
- [14] Mariana Gomide Panosso, Sílvia Regina de Souza, and Verônica Bender Haydu. 2015. Características atribuídas a jogos educativos: uma interpretação Analítico-Comportamental. *Revista Brasileira de Psicologia escolar e educacional* 19, 2 (2015), 233–241.
- [15] Francisco Ittalo Ribeiro Pessoa, Ana Liz Souto O Araujo, Wilkerson Andrade, and Dalton Guerrero. 2017. T-mind: um Aplicativo Gamificado para Estímulo ao Desenvolvimento de Habilidades do Pensamento Computacional. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, Vol. 28. 645.
- [16] R. Pressman and B Maxim. 2016. *Engenharia de software: uma abordagem profissional*. AMGH, Porto Alegre.
- [17] Gustavo Ramos, Laura Lessa, and Diana Cabral Cavalcanti. 2017. Disease Extermination: A Digital Educational Game on Vaccines, Viruses and Bacteria. In *Proceedings of the 23rd Brazilian Symposium on Multimedia and the Web*. ACM, 417–420.
- [18] Denise Maciel Sena, Elaine Harada T de Oliveira, and Leandro SG de Carvalho. 2014. Aplicativos móveis para o aprendizado de matemática. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, Vol. 25. 174.
- [19] Lotfi A Zadeh. 1965. Fuzzy sets. *Information and control* 8, 3 (1965), 338–353.

³A Lógica Fuzzy apresenta uma quantidade infinita de possibilidades entre os dois extremos, diferente da Lógica Booleana que permite apenas dois valores (0,1). Foi estruturada por [19] e faz parte da área de inteligência Artificial, muito utilizado para mensurar incertezas.

Assessment of the Treatment of Grammatical Aspects of Machine Translators to Libras

Luana Silva Reis
Universidade Federal da Paraíba
João Pessoa - Paraíba
luanareis@lavid.ufpb.br

Tiago Maritan U. de Araújo
Universidade Federal da Paraíba
João Pessoa - Paraíba
tiagomaritan@lavid.ufpb.br

Yuska Paola Costa Aguiar
Universidade Federal da Paraíba
Rio Tinto - Paraíba
yuska@dcx.ufpb.br

Manuella Aschoff C. B. Lima
Universidade Federal da Paraíba
João Pessoa - Paraíba
manuella.lima@lavid.ufpb.br

Angelina S. da Silva Sales
Universidade Federal da Paraíba
João Pessoa - Paraíba
angelina.sales@lavid.ufpb.br

ABSTRACT

Currently, a set of technologies has been developed with the aim of reducing barriers to access to information for deaf people, such as machine tools for sign languages. However, these technologies have some limitations related to the difficulty of handling some specific grammatical aspects of the sign languages, which can make the translations less fluent, and influence the deaf users experience. To address this problem, this study analyzes the machine translation of contents from Brazilian Portuguese (Pt-br) into Brazilian Sign Language (Libras) performed by three machine translators: ProDeaf, HandTalk and VLibras. More specifically, we performed an experiment with some Brazilian human interpreters that evaluate the treatment of some specific grammatical aspects in these three applications. As a result, we observed a significant weakness in the evaluation regarding the adequacy treatment of homonymous words, denial adverbs and directional verbs in the translations performed by the applications, which indicates the need for these tools to improve in the treatment of these grammatical aspects.

KEYWORDS

Machine translation, Brazilian sign language (Libras), deaf user, accessibility, quality of translation

1 INTRODUÇÃO

Para promoção da acessibilidade a conteúdos digitais e não digitais, assim como para facilitar a comunicação entre pessoas surdas e ouvintes, alguns recursos de Tecnologia Assistiva (TA) vêm sendo desenvolvidos nos últimos anos. Como exemplos, têm-se os aplicativos de tradução automática para Línguas de Sinais (LS): HandTalk¹, ProDeaf², Rybená³, VLibras⁴ etc. Estes tem como proposta reduzir as barreiras de acesso à informação das pessoas surdas, permitindo

¹<https://www.handtalk.me>

²<http://www.prodeaf.net/>

³<http://portal.rybena.com.br/site-rybena>

⁴<http://vlibras.gov.br>

In: XV Workshop de Trabalhos de Iniciação Científica (WTIC 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pósteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.

© 2018 SBC – Sociedade Brasileira de Computação.

ISBN 978-85-7669-435-9.

a visualização de conteúdos em LS na *web*, em dispositivos móveis, na TV, no cinema, entre outros.

Sendo a Libras uma língua genuína e independente da Língua Portuguesa, com estrutura gramatical própria [6, 14], no processo automático de tradução entre a língua-fonte e a língua-alvo pode ocorrer perda de informação [7]. Logo, a qualidade da tradução automatizada tem dividido opiniões entre os professores e intérpretes de Libras [4]. Em razão disso, é necessário que a tradução gerada possibilite ao usuário surdo a melhor correteza, compreensão e qualidade possível da sentença, minimizando a perda de informações, no processo de tradução.

Desta forma, considerando a relevância de assegurar a qualidade da tradução gerada por aplicativos desta natureza, neste trabalho estão descritos os procedimentos e resultados de um processo de avaliação realizado com os três tradutores automáticos do Português Brasileiro (Pt-br) para a Libras mais populares na loja *online* da plataforma *Android*: HandTalk, ProDeaf e VLibras. A proposta é investigar a correteza, compreensão e qualidade da tradução com relação ao tratamento de alguns aspectos gramaticais da Libras, tais como: palavras homônimas, advérbios de negação e verbos direcionais. A escolha dos aspectos gramaticais a serem avaliados deu-se pela seleção aleatória, a partir dos 11 aspectos gramaticais identificados nesse trabalho, cujos os tratamentos inadequados podem interferir na qualidade das sentenças geradas, a saber: hominímia, advérbios (negação, modo e intensidade), verbos (direcionais, que concordam com a localização e classificadores), regionalismo, *crossover* e frases (interrogativas e exclamativas). Os testes foram conduzidos com 16 intérpretes de Libras que traduziram 12 sentenças para Libras e posteriormente avaliaram as sentenças equivalentes, geradas automaticamente pelas três aplicações mencionadas.

O artigo está organizado em 5 seções. Na Seção 2 são apresentados os trabalhos relacionados à verificação da qualidade da tradução gerada por tradutores automáticos para LS. Na Seção 3, descreve-se a metodologia adotada para a execução das oficinas de avaliação dos tradutores automáticos por intérpretes de Libras. Na Seção 4, os resultados obtidos com o estudo são socializados e discutidos. Por fim, na Seção 5 têm-se as considerações finais.

2 TRABALHOS RELACIONADOS

Na literatura científica existem trabalhos voltados para avaliação da tradução automática de sentenças do Pt-br para a Libras.

No estudo de Colling e Boscaroli [3] a avaliação realizada considerou a tradução de frases pelos tradutores HandTalk, ProDeaf e Rybená em contexto interpretativo pelo viés da criança surda. Os autores identificaram que as crianças tiveram dificuldades em compreender as traduções geradas pelos *softwares*.

No trabalho de Soares et al. [15] a utilização de uma solução de tradução automática de qualquer língua oral estrangeira para Libras, utilizando o aplicativo VLibras, foi avaliada com usuários surdos de Libras. A pesquisa mostra que os usuários conseguiram responder com precisão média de 40%, as perguntas realizadas sobre as traduções obtidas.

Vieira et al. [16] consideraram o tratamento de “Expressões Não-Manuais” a partir de traduções dos aplicativos HandTalk e ProDeaf, sob a perspectiva de usuários surdos. Como resultado, os participantes indicaram que os avatares não sinalizavam expressões não-manuais, dificultando a compreensão das traduções geradas.

No estudo de Lima [11] foram identificados quinze problemas relacionados ao tratamento de contexto no processo da tradução do Pt-br para Libras para a Suíte VLibras. Uma nova versão do componente de tradução foi integrado a Suíte e na avaliação posterior ao refinamento verificou-se melhor compreensão dos conteúdos pelos usuários surdos.

Embora a qualidade das traduções automáticas tenha sido investigada nos trabalhos supracitados, nenhum destes considerou a análise específica da “Adequação de Palavras Homônimas”, “Tratamento de Advérbios de Negação” e “Tratamento de Verbos Direcionais” pela visão de intérpretes Libras. Uma vez que, o profissional intérprete de Libras é especialista em ambas as línguas (Pt-br e Libras), torna-se indispensável a sua participação na avaliação da tradução gerada por tradutores automáticos. Além disso, é importante destacar que o atual estudo contempla as três aplicações com maior número de downloads, o que permite inferir que a avaliação realizada atinge uma quantidade expressiva de usuários.

3 METODOLOGIA

A metodologia adotada considerou três etapas para o estudo: Planejamento, Condução e Análise dos Dados.

Planejamento: identificação dos aspectos gramaticais de Libras cujos os tratamentos inadequados podem interferir na qualidade das sentenças geradas, a saber: homonímia, advérbios (negação, modo e intensidade), verbos (direcionais, que concordam com a localização e classificadores), regionalismo, *crossover* e frases (interrogativas e exclamativas). Considerando a complexidade para contemplar todos os aspectos mencionados em apenas um experimento e a inviabilidade diante do tempo disponível, se fez necessário delimitar o escopo do experimento. Optou-se por selecionar, aleatoriamente, três aspectos gramaticais: homonímia, advérbios de negação e verbos direcionais.

A Homonímia é presente quando duas ou mais palavras de origens e significados distintos têm a mesma grafia e/ou fonologia [10]. As palavras homônimas e as respectivas sentenças contempladas no estudo foram extraídas, de forma aleatória, do Dicionário Enciclopédico Ilustrado Trilíngue da Libras [2]. As palavras sorteadas e as respectivas sentenças contendo significados (ou sinais) diferentes foram: (i) MANGA: A **manga** da camisa está rasgada (parte de roupa) / A **manga** é uma fruta típica do verão (fruta); (ii) ESCALAR:

O menino **escalou** o muro para fugir do cachorro (subir) / O diretor do hospital **escalou** o médico para o plantão (escolher); e, (iii) COMBATER: A ação do governo não conseguiu **combater** às drogas (extinguir) / Os países decidiram dar uma trégua no **combate** (guerrear).

Os Advérbios de Negação podem ser representados de três maneiras em Libras [8]: acréscimo do sinal NÃO à frase afirmativa; a incorporação de um movimento contrário ou diferente ao do sinal negado; aceno de cabeça (simultâneo) com a ação que está sendo negada. Os três verbos de negação e as respectivas sentenças contempladas no estudo foram selecionados, aleatoriamente, de [8] e do Dicionário Enciclopédico Ilustrado Trilíngue da Libras [2], respectivamente. A saber: (i) Acréscimo do sinal NÃO à frase afirmativa: “Não adianta gritar, ele **não ouve**”; (ii) Incorporação de um movimento contrário ou diferente ao do sinal negado: “**Não gosto** desse seu jeito de agir”; e, (iii) Aceno de cabeça (simultâneo) com a ação que está sendo negada ou juntamente com as outras formas de negação: “Eu **não posso** atendê-lo”.

Os Verbos Direcionais, na Libras, são aqueles que flexionam-se em pessoa, número e aspecto [5]. Para a avaliação pretendida três verbos direcionais e as respectivas sentenças foram selecionados aleatoriamente com base no estudo descrito por Quadros [5] e Felipe [8]. A saber: (i) **PERGUNTAR**: Eu pergunto para você. [8]; (ii) **DIZER**: Eu disse a vocês. [5]; e, (iii) **DAR**: Ele deu para mim. [5].

A seleção das aplicações de tradução automática de Pt-br para Libras tomou como base: (i) a gratuidade do acesso à aplicação, (ii) a possibilidade de uso dos aplicativos em dispositivos móveis com a plataforma *Android*⁵, (iii) possuir mais de 50.000 *downloads* e avaliação superior à 4.0 na Play Store. Em razão disso, foram selecionados os aplicativos HandTalk (+1.000.000 de *downloads* e avaliação média de 4,6), ProDeaf (+1.000.000 de *downloads* e avaliação média de 4,6) e VLibras (+50.000 de *downloads* e avaliação média de 4,2).

O recrutamento de participantes considerou profissionais intérpretes com experiência nas duas línguas (Pt-br e Libras). Esta escolha permite uma análise comparativa da tradução feita pelos intérpretes com aquelas geradas de forma automática pelas aplicações. As avaliações foram realizados em 6 instituições de ensino (Municipal, Estadual e/ou Federal) do estado da Paraíba e contou com o engajamento de 16 (dezesseis) intérpretes de Libras com idades entre 25 e 40 anos, todos com pelo menos 5 anos de experiência como intérpretes.

Para apoiar a coleta de dados, foi elaborado um questionário composto por 3 seções, sendo cada uma dedicada as sentenças relativa a um dos aspectos gramaticais de Libras investigados neste estudo: palavras homônimas, advérbios de negação e verbos direcionais. O questionário totaliza doze itens a serem avaliados sobre (i) corretude gramatical da tradução; (ii) compreensão sobre a tradução gerada e, (iii) qualidade geral da tradução. Cada item dispõe de uma escala de seis pontos (1=péssimo à 6=perfeito) para avaliação da qualidade da sentença. Este tipo de escala foi adotada por evitar avaliações majoritariamente neutras [12] e por ter sido utilizadas em trabalhos da mesma natureza [1, 13].

⁵De acordo com Gomes et. al [9], é a plataforma para dispositivos móveis mais popular do mundo.

Condução: A preparação do ambiente de avaliação consistiu no posicionamento de câmeras para captura de áudio e vídeo, seguida da assinatura de termo de consentimento⁶ por parte dos intérpretes participantes, que foram previamente instruídos sobre os procedimentos da avaliação.

O experimento foi conduzido de forma individual e não paralela para cada participante. Cada participante foi instruído a fazer a sinalização para todas as sentenças contempladas no estudo, sendo estas registradas em vídeo para posterior análise comparativa aquelas geradas pelas aplicações. Além disso, os avaliadores observaram os intérpretes neste processo para tomar nota dos aspectos relevantes para análise.

Em seguida, cada intérprete respondeu um questionário de avaliação. Na ocasião, eles tiveram acesso aos vídeos gerados pelas aplicações para as 12 sentenças. A avaliação foi feita de forma cruzada⁷ para permitir que cada intérprete de Libras avaliasse as três aplicações de tradução automática e evitar que o teste fosse muito extenso, demorado e cansativo, o que poderia interferir no resultado. Com isso, evitaria-se também que a ordem de apresentação das aplicações influenciasse nos resultados do teste.

Análise dos Dados: O processo de análise dos dados obtidos compreendeu os registros de áudio e vídeo feitos pelos intérpretes na tradução das sentenças, nas anotações dos avaliadores durante a observação dos participantes e nas respostas obtidas via questionário. A análise foi dirigida à identificação das semelhanças e discrepâncias entre a tradução realizadas pelos intérpretes (humanos) e a traduções geradas pelas aplicações (automáticas).

4 RESULTADOS DA AVALIAÇÃO

Todas as seções do experimento foram realizadas na íntegra, sem abandono dos intérpretes ou interrupção pelos avaliadores. A duração média foi de 15 minutos, incluindo a tradução das 12 sentenças pelo intérprete (registro em áudio e vídeo), seguida da avaliação, via questionário, das traduções automáticas geradas pelas aplicações. Os principais resultados encontram-se na Tab. 1 e Fig. 1.

Para as **Palavras Homônimas** consideradas no estudo (MANGA, ESCALAR, COMBATER), das seis sentenças trabalhadas, apenas a aplicação ProDeaf fez a distinção contextual adequada para o uso da palavra COMBATER. O tradutor HandTalk apresentou a sinalização correta para apenas um dos dois sentidos das palavras homônimas consideradas no estudo. Logo, não foi possível perceber o tratamento de contexto para as palavras avaliadas, uma vez que esta é atribuída à apenas um sinal, mesmo quando possui diferentes significados. No caso da palavra ESCALAR, os tradutores ProDeaf e VLibras apresentam dupla inadequação para as sentenças avaliadas, apresentando sinais incoerentes com o contexto das sentenças. No entanto, a média dos três aplicativos para a Corretude, Compreensão e Qualidade das sentenças permaneceu semelhante, como pode ser verificado na Fig. 1.

Para os **Advérbios de Negação** (NÃO OUVIR, NÃO GOSTAR e NÃO PODER) as aplicações apresentaram traduções heterogêneas. Os aplicativos ProDeaf e HandTalk apresentaram a incorporação adequada dos advérbios de negação para as três sentenças avaliadas, o que pode ser verificado na Fig. 1, com a semelhança das avaliações

desses dois aplicativos no que se refere a Corretude, Compreensão e Qualidade dessas traduções. O aplicativo VLibras apresentou o tratamento adequado apenas para o advérbio de negação “NÃO OUVIR”, apresentando assim um tratamento parcial desse aspecto.

Para os **Verbos Direcionais** (PERGUNTAR, DIZER, DAR), o tratamento realizado pelo tradutor HandTalk foi adequado e satisfatório para as três sentenças do estudo. Esta afirmativa se apoia na similaridade da tradução automática à tradução dos intérpretes, mas também na avaliação feita por estes via questionário. Esse resultado pode ser verificado na Fig. 1 onde é observada uma melhor avaliação nos resultados do aplicativo Handtalk em relação aos demais aplicativos. Por outro lado, as aplicações ProDeaf e VLibras além de não incorporarem os aspectos de tratamento para a direcionalidade de verbos para os exemplos testados, forneceram português sinalizado para as sentenças e não Libras.

Diante do exposto é possível identificar que, de forma geral, os três tradutores avaliados apresentam algumas adequações e inadequações para o tratamento contextual dos aspectos gramaticais analisados. A respeito das palavras homônimas avaliadas, apenas o tradutor HandTalk fez distinção de sinal na tradução das duas sentenças avaliadas para a palavra "Combater". Para o tratamento de Advérbios de negação os aplicativos ProDeaf e HandTalk apresentaram corretamente a incorporação da negação para as três sentenças avaliadas, enquanto que o aplicativo VLibras, apresentou a incorporação correta apenas para a negação da palavra "Ouvir". Em relação a direcionalidade do verbo, o aplicativo HandTalk apresentou a incorporação desse aspecto de forma satisfatória para as três sentenças avaliadas, enquanto que para o esse mesmo aspecto gramatical, o aplicativo ProDeaf incorporou corretamente a direcionalidade apenas para a sentença da palavra "Perguntar".

Um resumo dos resultados da adequação e inadequação dos aspectos gramaticais avaliados, pode ser visualizado na Tab. 1.

Tabela 1: Resultados da Adequação dos Aspectos Gramaticais Avaliados.

	Termos	HandTalk	ProDeaf	VLibras
S1	Manga(parte de roupa)	Não	Não	Não
	Manga (fruta)	Sim	Sim	Sim
	Escalar (subir)	Não	Não	Não
	Escalar (escolher)	Não	Não	Não
	Combater (extinguir)	Não	Sim	Sim
	Combater (guerrear)	Sim	Sim	Não
S2	Não ouvir	Sim	Sim	Sim
	Não gostar	Sim	Sim	Não
	Não poder	Sim	Sim	Não
S3	Perguntar	Sim	Sim	Não
	Dizer	Sim	Não	Não
	Dar	Sim	Não	Não

5 CONSIDERAÇÕES FINAIS

O processo de tradução automática entre Pt-br para Libras é complexo e não trivial, uma vez que as estruturas gramaticais das línguas envolvidas não são equivalentes. Neste sentido, diferentes aspectos gramaticais precisam ser considerados para minimizar a perda de

⁶Disponível em: <https://tinyurl.com/termo-de-autorizacao>

⁷Esquema do teste cruzado disponível em: <https://tinyurl.com/questionario-cruzado>

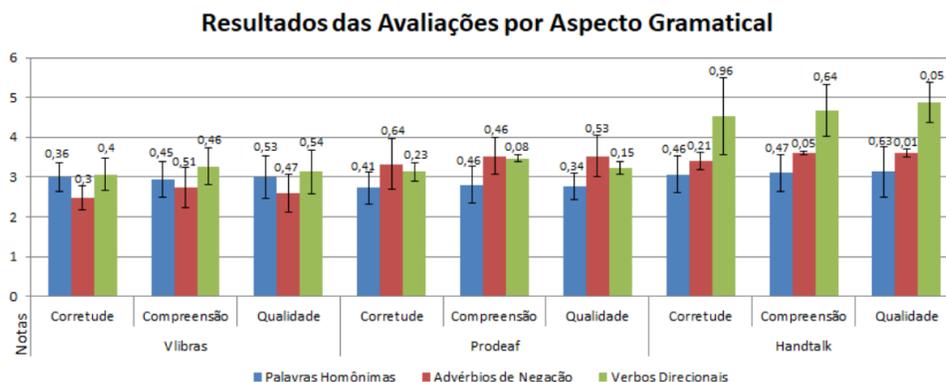


Figura 1: Média e Desvio Padrão das Avaliações dos Aspectos Gramaticais por Tradutor Automático utilizando uma escala de seis pontos (1=péssimo à 6=perfeito).

informação e favorecer a compreensão do conteúdo que se deseja comunicar.

Nesta pesquisa, foram avaliados os aspectos gramaticais: Palavras Homônimas, Advérbios de Negação e Verbos Direcionais com a utilização de um teste cruzado, contemplando a participação de 16 intérpretes, para avaliar estes aspectos para um conjunto de 12 sentenças traduzidas pelas aplicações HandTalk, ProDeaf e VLibras.

Os resultados alcançados nos permitem afirmar que, para os exemplos testados, o tratamento dos Advérbios de Negação foi adequado e equivalente às traduções dos intérpretes, para todas as traduções automáticas disponibilizadas nos aplicativos HandTalk e ProDeaf e parcialmente adequado pelo aplicativo VLibras. O tratamento dos Verbos direcionais foi avaliado como adequado apenas para a aplicação HandTalk, uma vez que as aplicações ProDeaf e VLibras forneceram a sinalização em datilografia e não em Libras, nos exemplos testados. Para o tratamento das Palavras Homônimas, diante da quantidade de termos considerados no estudo, generalizações não podem ser feitas sobre os resultados. Como apenas a aplicação ProDeaf fez distinção contextual para apenas um dos três termos contemplados no estudo, pode-se inferir que este aspecto gramatical está presente no tradutor. No entanto, não é possível indicar a abrangência da solução para um conjunto mais amplo de palavras homônimas.

Como soluções para a adequação de palavras homônimas, sugere-se a utilização de um número maior de sinais para os diversos significados de uma palavra homônima e a avaliação do contexto da sentença. Para a adequação de advérbios de negação é necessário a avaliação das três formas de representação dos advérbios de negação na Libras e com qual dessas três formas o verbo deve ser representado, de acordo com suas particularidades. Para adequação de verbos direcionais é necessário a implementação da direcionalidade dos verbos na Libras, para que no processo de geração da tradução o movimento do avatar varie de acordo com a posição das pessoas que estão envolvidas no diálogo.

Como proposta de trabalhos futuros, tenciona-se um estudo mais aprofundado nos aspectos gramaticais avaliados nessa pesquisa, com a participação de usuários surdos. Além disso, a realização da avaliação do tratamento dos outros aspectos gramaticais identificados durante a execução deste trabalho: advérbios de intensidade e

de modo, verbos que concordam com a localização e classificadores, regionalismos, crossover, e frases interrogativas e exclamativas.

REFERÊNCIAS

- [1] Tiago Maritan Ugolino de Araújo. 2012. Uma solução para geração automática de trilhas em língua brasileira de sinais em conteúdos multimídia. (2012).
- [2] Fernando César Capovilla. 2001. *Dicionário enciclopédico ilustrado trilíngüe da língua de sinais brasileira: sinais de M a Z*. Vol. 2. Edusp.
- [3] João Paulo Colling and Clodis Boscardioli. 2014. Avaliação de tecnologias de tradução Português-Libras visando o uso no ensino de crianças surdas. *RENOTE* 12, 2 (2014).
- [4] Ygor Corrêa, Rafael Peduzzi Gomes, and Vinicius Gadis Ribeiro. Aplicativos de Tradução Português-Libras na Educação Bilíngüe: desafios frente à desambiguação. *RENOTE* 15, 2 (????).
- [5] Ronice Müller de Quadros. 2009. *Educação de surdos: a aquisição da linguagem*. Artmed Editora.
- [6] Ronice Müller de Quadros and Lodenir Becker Karnopp. 2009. *Língua de sinais brasileira: estudos lingüísticos*. Artmed Editora.
- [7] Umberto Eco. 2007. *Quase a mesma coisa*. Editora Record.
- [8] Tanya A Felipe. 2007. *Libras em contexto: curso básico: livro do estudante*. (2007).
- [9] Rafael Caveari Gomes and Jean Alves Fernandes. 2012. Sistema Operacional Android. *Universidade Federal Fluminense* (2012).
- [10] Es Cruz da Lima. Alguns aspectos semânticos da LIBRAS: Um estudo do léxico de seus sinais em suas relações de Sinonímia, Homonímia, Homógrafas e Polissemia. In *Congreso Internacional Aasociació*.
- [11] Manuella A.C.B. Lima, Tiago M.U. de Araújo, and Erickson S. de Oliveira. 2015. Incorporation of Syntactic-Semantic Aspects in a LIBRAS Machine Translation Service to Multimedia Platforms. In *Proceedings of the 21st Brazilian Symposium on Multimedia and the Web (WebMedia '15)*. ACM, New York, NY, USA, 133–140. <https://doi.org/10.1145/2820426.2820434>
- [12] Sara Morrissey. 2008. *Data-driven machine translation for sign languages*. Ph.D. Dissertation. Dublin City University.
- [13] Rubén San-Segundo, Juan Manuel Montero, R Córdoba, Valentin Sama, F Fernández, LF D'Haro, Verónica López-Ludeña, D Sánchez, and Antonio García. 2012. Design, development and field evaluation of a Spanish into sign language translation system. *Pattern Analysis and Applications* 15, 2 (2012), 203–224.
- [14] Ana Paula Santana. 2007. *Surdez e linguagem*. Grupo Editorial Summus.
- [15] Renan Soares, Rostand Costa, Manuella Aschoff, and Tiago Maritan. 2017. AnyLanguage-To-LIBRAS: Evaluation of an Machine Translation Service of Any Oralized Language for the Brazilian Sign Language. In *Proceedings of the 23rd Brazilian Symposium on Multimedia and the Web (WebMedia '17)*. ACM, 481–488.
- [16] Maristela C Vieira, Ygor Corrêa, Lucila MC Santarosa, and MCV BIAZUS. 2014. Análise de expressões não-manuais em avatares tradutores de Língua Portuguesa para Libras. In *XIX Conferência Internacional sobre Informática na Educação-TISE*.

Reformulação do Projeto de Usabilidade da Ferramenta Lagarto

Samir Chaves

Universidade Federal do Ceará – UFC
Campus do Pici, s/n
Fortaleza, Ceará CEP 60451-970
samirchaves@alu.ufc.br

Fernando Trinta

Universidade Federal do Ceará – UFC
Campus do Pici, s/n
Fortaleza, Ceará CEP 60451-970
fernando.trinta@dc.ufc.br

ABSTRACT

The development of Location Based Mobile Games (LBMG) is a complex task. To overcome this problem, authoring tools have been proposed aiming at reducing the time and effort required to create these applications. One of these tools is LAGARTO, a web authoring application to help non-programmers users to create multiplayer LBMGs with augmented reality features. The first version of LAGARTO presented a usability level below end-users expectations. Therefore, this research proposes a new usability project for the LAGARTO editor to address its current problems. This document presents the results already obtained from a work in progress, with promising results of improvement in the usability of LAGARTO.

KEYWORDS

Location-Based Games, Game Modelling, Authoring Tools

1 INTRODUÇÃO

A nova era de entretenimento e interação dos jogos pervasivos, traz à tona novas possibilidades de conversação entre a tecnologia e o mundo real na experiência dos usuários, porém com uma ainda complexa metodologia de desenvolvimento. Sucessos de títulos como Pokemon Go e Ingress mostraram o potencial atrativo que jogos baseados em localização e com recursos de realidade aumentada podem trazer para diversas áreas, como por exemplo, turismo ou educação [3] [5].

Porém, a criação destas aplicações ainda demanda um grande esforço de diversos profissionais envolvidos em diversas tarefas de seu processo de desenvolvimento, como projeto do jogo, programação e testes. Neste sentido, ferramentas de autoria assumem um papel importante, uma vez que seu objetivo está em facilitar a criação de aplicações. Uma dessas ferramentas é a LAGARTO[4], uma proposta acadêmica para autoria de Jogos Móveis Baseados em Localização (JMBL), desenvolvida na Universidade Federal do Ceará como projeto de mestrado. A LAGARTO tem como objetivo maior permitir que pessoas sem conhecimento prévio de programação ou habilidades de desenvolvimento de aplicativos possam projetar e implementar JMBLs de forma facilitada.

No projeto original da LAGARTO, a interação de autores com a ferramenta é feito por meio da aplicação web chamada Editor, responsável tanto pela criação quanto pelo gerenciamento dos jogos planejados. O editor original propôs uma notação visual e uso de certas tecnologias que não permitiram uma boa usabilidade

aos usuários da ferramenta, principalmente devido ao uso *plugins* pré-estabelecidos, que impuseram um modelo de diagramação não intuitivo aos usuários. Tendo em vista este problema, foi elaborado um projeto de pesquisa de iniciação científica que se propôs a estudar e propor uma nova interface e novos modos de interação para designers de jogos com a LAGARTO.

Este documento apresenta os resultados obtidos até agora no projeto. Para isso, na seção seguinte é apresentada uma visão geral da LAGARTO. Na seção 3 é reportada a metodologia aplicada neste projeto de pesquisa. A quarta seção apresenta os resultados até obtidos com a reformulação da interface do editor da LAGARTO. Por fim, a última seção conclui o documento e apresenta próximos passos desta pesquisa.

2 LAGARTO

LAGARTO (*LocAction Games AuthoRing Tool*) [4] é uma ferramenta de autoria para JMBLs que busca jogos móveis baseados em localização que busca abstrair a complexidade de diversos aspectos do desenvolvimento de softwares móveis e sensíveis à localização. A ferramenta faz com que o usuário exerça a função de *game designer*, elaborando a lógica do jogo de forma visual por meio de um editor web. A estrutura deste jogo é recuperada e executada em *smartphone*, por meio de um aplicativo aplicativo interpretador. Além de permitir a autoria dos jogos, a LAGARTO também gerencia o ciclo de vida dos jogos em execução. Este ciclo inclui a disponibilização do jogos para os jogadores, atualização de jogos e interrupção de funcionamento do servidor.

Os jogos projetados na LAGARTO seguem um modelo de aplicação definido por [2], onde cada jogo é visto como um conjunto de missões a serem realizadas por um jogador ou grupo de jogadores. Cada missão é composta por uma ou mais mecânicas comuns à JMBLs[1], como deslocamento à um certo ponto do mapa, captura de objetos virtuais, dentre outras ações. Missões podem ser compartilhadas entre grupos de usuários, criando a ideia de colaboração entre equipes. Também é possível estabelecer prioridades entre missões e mecânicas, indicando se determinadas ações precisam ser realizadas antes que outras. Na sua versão inicial, a interface da ferramenta foi criada conforme figura 1.

Apesar de atender a criação dos jogos de forma visual, a interface de interação proposta para primeira versão da LAGARTO apresentou problemas. Os resultados obtidos a partir de testes de usabilidade revelaram que a LAGARTO possui apenas usabilidade considerada mediana em relação à outros softwares, e que alguns dos recursos da linguagem ainda confundem ou não são disponibilizados intuitivamente para os usuários. Os testes realizados com usuários não programadores foram essenciais para realizar o diagnóstico dos problemas de usabilidade, onde a busca por soluções a estes problemas é de fundamental importância para a aceitação

In: XV Workshop de Trabalhos de Iniciação Científica (WTIC 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pósteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.

© 2018 SBC – Sociedade Brasileira de Computação.
ISBN 978-85-7669-435-9.

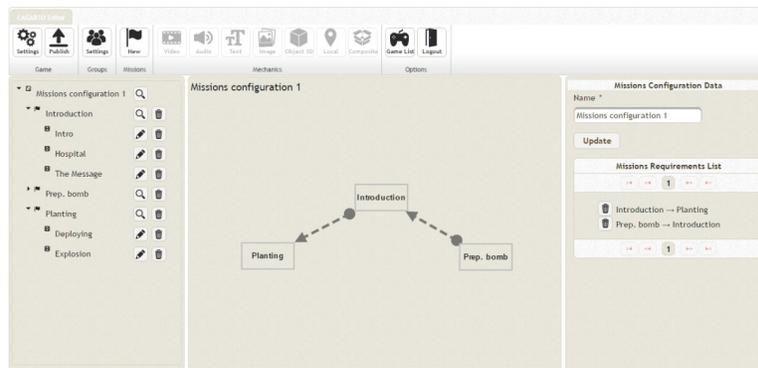


Figura 1: Componentes do Editor - LAGARTO.

e popularização da ferramenta. Dentre os problemas citados nos experimentos podemos destacar (i) a insegurança na realização de ações como alternar entre mecânicas e missões, (ii) dúvidas em relação ao significado de alguns ícones e (iii) confusão quando da implementação de recursos específicos relacionados às mídias suportadas pela ferramenta.

A melhor forma de solucionar estas questões parte pela realização de um estudo de usabilidade mais detalhado, visando implementar uma reformulação da interface gráfica da LAGARTO. Espera-se que a inclusão de recursos como arrastar-soltar (do inglês, *Drag and Drop*), a reorganização das funções da ferramenta e a remodelagem de ícones proporcionarão maior facilidade de uso da ferramenta. Consequentemente, estas mudanças farão com que os usuários sintam-se mais confiantes durante o desenvolvimento dos jogos, gerando benefícios a todos os usuários, pesquisadores e professores que utilizam a LAGARTO.

3 METODOLOGIA DA PESQUISA

A partir da detecção dos problemas de usabilidade da LAGARTO, foi elaborado um projeto de pesquisa de iniciação científica com objetivo geral de definir e implementar um novo projeto da interface gráfica do editor da LAGARTO. Para isso, o projeto (ainda em andamento) prevê a realização de cinco etapas: (i) Estudos sobre ferramentas de autoria, (ii) Estudo sobre a primeira versão da LAGARTO, (iii) Elaboração do projeto da nova interface do Editor, (iv) Implementação da nova interface em uma tecnologia adequada e (v) Testes com usuários finais. Atualmente, o projeto encontra-se na quarta etapa, com previsão de conclusão para final deste ano.

O primeiro passo foi realizado com a leitura do material existente (dissertação de mestrado e artigos científicos) sobre a própria LAGARTO, que traziam uma série de ferramentas de autoria e demais conceitos relacionados ao tema de jogos móveis baseados em localização. Destes trabalhos foram coletadas boa parte de trabalhos relacionados que serviram de base de estudos para características interessantes para a interface gráfica do Editor. Além disso, outros trabalhos também foram selecionados por meio de buscas na Internet. Dentre eles, destaca-se a TaleBlazer¹ do MIT, aplicação de objetivo bastante semelhante ao LAGARTO, porém com uma

implementação diferente dos conceitos que compõem um jogo. Os elementos inspirados no TaleBlazer são relacionados ao layout. Por exemplo, a todo instante um menu flutuante se encontra sobre a tela com opções gerais sobre o jogo. Esta abordagem foi tida como positiva, permitindo expandir a área de visualização do jogo. Além disso, uma grande área da janela de visualização é dedicada ao mapa que suporta o jogo, o que proporciona uma melhor usabilidade já que há um foco nessa região. Por fim, a inserção de um menu de contexto também foi bem avaliada. Este menu apresenta a lista de opções ao se clicar com o botão direito sobre a região do mapa, mostrando funções disponíveis relacionadas ao contexto daquele local.

Em relação à segunda etapa, o estudo sobre a proposta inicial da ferramenta indicou que o layout existente no Editor contava com tema relativamente rústico, efeito trazido tanto por elementos como fontes e ícones escolhidos, como pela aplicação, em diversos momentos, de texturas à certas áreas. Atualmente, essa abordagem não é tida como mais adequada para aplicações semelhantes à LAGARTO. Essa informação passada visualmente, por mais subjetiva que seja, leva a conquista ou não de novos usuários, sendo assim de suma importância para a popularidade da aplicação. Além disso, a organização dos elementos da página tornava o foco central do usuário destoante, comprometendo assim a usabilidade.

A maior parte da navegação baseava-se em painéis fixos que ocupavam uma parcela considerável da janela de trabalho do editor. Este recurso foi tido como um problema, pois não deixava claro, pela quantidade de opções disponíveis, qual o passo seguinte a ser tomado em meio a criação ou edição de um jogo. Outro ponto importante problemático foi em relação à modelagem do jogo, que era realizada em uma área acessível pelos painéis. Como um jogo é constituído de missões, cada missão possui sua própria modelagem, o que obriga a controlar qual missão aparece por vez na tela, trazendo um esforço maior quando o objetivo é ter um panorama geral da organização do jogo. Ainda na questão da organização das missões, estas são formadas por mecânicas, que por sua vez representam diretamente locais geográficos reais. O fato das mecânicas serem vistas em um painel opaco, sem informação palpável de onde se localizam, também abre margem a melhorias. Por fim, um problema também identificado foi o uso exagerado de janelas

¹<http://www.taleblazer.org/>

modais para interação do usuário. Estas janelas não são um recurso tão utilizados em projetos mais modernos de interfaces web.

Atualmente, esta pesquisa se encontra nas terceira e quarta etapas do projeto originalmente proposto. A partir de um processo de implementação iterativo e incremental, uma nova interface para o Editor da LAGARTO vem sendo construído, com ideias de melhoria sendo propostas à cada nova iteração. No início do processo de implementação, um rascunho com a proposta para a nova interface do usuário foi proposta. Esta interface foi inicialmente prototipada em papel para estabelecimento de princípios básicos a serem seguidos. Estes princípios incluíram manter um visual limpo da ferramenta, fazer uso extensivo de ações *drag-and-drop* e maximizar a visão do mapa do jogo pelo usuário. A seção a seguir detalha melhor os resultados até obtidos nesta etapa.

4 EDITOR 2.0 - LAGARTO

Atualmente, o Editor está sendo redesenhado de uma forma a preservar a essência e funcionalidades principais da versão original, contudo, trazendo um novo esquema de usabilidade. Essas modificações prezam por compor uma maior dinamicidade ao manejo, por meio de uma interface moderna e livre de elementos poluentes e distrativos. A Figura 2 apresenta a nova proposta da Ferramenta.

Após o desenho de alguns esboços do layout da ferramenta, foi escolhido um a ser implementado. Este prezava pela mesma ideia do Google Maps, com o elemento mapa ocupando praticamente toda a tela, porém com um sutil menu lateral. Além disso, é observado a presença de painéis arrastáveis pelos quais serão configurados os elementos dos jogos. Também existe um botão com a função de adicionar componentes dos jogos (times, missões e mecânicas) ao contexto em questão. A personalização trazida pelo posicionamento da janela, aliado à simplicidade dos elementos, e a visão panorâmica do mapa compõem a ambientação desejada. Ainda se tratando do esboço selecionado, nele já é possível notar a estilização trazida aos componentes. A iconografia com um traço moderno, o uso de leves sombras para a sensação de sobreposição, a organização de espaçamento e alinhamento dos objetos, entre outras coisas, carregam juntas a configuração da proposta visual mais limpa.

A nova versão do editor está sendo implementada com uso intensivo de ReactJS e da GoogleMaps API. O ReactJS é um framework Javascript voltado para criação de interfaces web criado pelo Facebook. A sua escolha foi graças (i) sua grande aceitação pela comunidade de desenvolvedores, (ii) sua rápida curva de aprendizado, (iii) seu ambiente de desenvolvimento simples e prático, e por, principalmente, (iv) ter a propriedade de tratar componentes web isoladamente, o que gera uma organização instantânea do código fonte e desenvolvimento mais rápido e robusto. Outra questão importante é a facilidade de integrar-se com o padrão Redux, que permite armazenar um estado global da aplicação e emitir eventos de alteração deste estado, o que traz uma montagem ideal para um webapp. É pretendido que este estado seja posteriormente usado pelo servidor da aplicação para seu armazenamento em banco de dados. Por tornar o código mais reusável, o ReactJS permitiu a mais prática criação dos componentes criados pelo esboço.

Já a GoogleMaps API permite a interação do usuário a todo o tempo com um mapa, definindo posições reais de seu interesse, e modelando toda a dinâmica de seu jogo sobre o local escolhido para

suportá-lo. Esta API é majoritariamente utilizada no projeto, desde a criação do mapa até a função de posicionamento dos marcadores das mecânicas. O objetivo é que praticamente toda a gestão do jogo seja realizada no próprio mapa. Para colocar isso em prática, alguns objetos foram criados. O primeiro deles é o *Marker*, o qual representa uma mecânica no modelo de aplicação da LAGARTO. Este se encontra no local do mapa onde a mecânica deve ser realizada e pode ser reposicionado a qualquer momento. A modelagem de prioridade e ordem de execução das mecânicas dentro de cada missão é feita através de arestas que conectam os *markers*.

Como a API não oferece uma estrutura de dados suficiente para atender a esse propósito, foi criada uma própria. Uma inovação em relação à interface original da LAGARTO, é que foi incorporada uma animação à aresta indicando que a partir de um *marker*, qual o próximo local a ser visitado. Esta era uma das principais queixas dos usuários que participaram do experimentos com a versão original do editor, pois a forma anterior não deixava claro qual mecânica ou missão precedia a outra. Cada aresta de uma missão é representada por uma *Polyline*, o segundo objeto a ser usado. O conjunto de *polylines* e mecânicas pertencentes a uma missão são formados por uma estrutura de dados na forma de um Grafo Direcionado Ponderado, permitindo armazenar a posição das mecânicas, as suas prioridades e a ordem de sua execução. O recurso de *context menu* também é utilizado para fornecer opções dos marcadores e arestas de maneira mais rápida.

4.1 Componentes do Editor

O nova versão do editor da LAGARTO possui 5 componentes principais, pelos quais, além do mapa, se dará toda a edição do jogo. As subseções a seguir descrevem cada um destes elementos.

4.1.1 Lista de Times. Na hierarquia dos componentes de um jogo, um time é o de nível mais alto, pois este é constituído por missões que por sua vez possuem mecânicas. O painel que os lista é um painel de posicionamento fixo responsável pela criação de novos times e, além disso, sua seleção para modificação das respectivas propriedades. Cada time criado possui uma cor de identificação, a qual está presente também nos marcadores de mecânicas que o compõem, com o intuito de facilitar sua identificação em meio ao mapa. Posteriormente, pretende-se adicionar funcionalidades de remoção de times e renomeio ao duplo clique a esse componente.

4.1.2 Lista de Missões. Neste menu móvel é possível visualizar toda a hierarquia do jogo. Por meio de listas retráteis, têm-se, respectivamente, os times, as missões e mecânicas organizadas de forma a tornar a árvore hierárquica intuitiva. Nele também é possível selecionar cada um desses elementos. Ao selecionar um time, há duas modificações aparentes na visão do usuário: todas as missões presentes no mapa que não o pertencem aumentam sua transparência, destacando assim as que o pertencem, e o painel de propriedades mostra as informações a ele referentes. Ao selecionar uma missão, acontece coisa semelhante: todas as mecânicas que não a constituem diminuem sua opacidade e o painel de propriedades também mostra as respectivas informações. O mesmo acontece ao selecionar uma mecânica. Em breve, almeja-se adicionar uma barra de ferramentas a esse menu, com as opções de “Criar Novo”, “Remover” e “Duplicar”, válidas para qualquer elemento selecionado.



Figura 2: Componentes do Editor - LAGARTO.

Além disso, é pretendido que seja possível mover uma missão de um time para outro apenas por meio de um *drag and drop*.

4.1.3 Painel de Propriedades. Ao selecionar um, e apenas um, item do jogo, o painel de propriedade mudará seu conteúdo de acordo com o tipo do componente selecionado, contendo suas configurações particulares. Atualmente, a única informação comum a todos é o nome descritivo. No caso de um time, suas propriedades são: nome, cor e lista de missões que o compõem. Em uma missão estão presentes o nome, o time a qual pertence e as mecânicas que a constituem. No caso da mecânica, apresentam-se nome, missão a qual pertence, tipo e localização. Ao se modificar uma informação, esta mudança é populada automaticamente em toda a aplicação tanto em dados armazenados em *cache* quanto na própria visão do usuário, permitindo ter um compreensão imediata do acabou de editar.

4.1.4 Menu de Adição de Mecânicas. Estando a todo momento sobre o mapa, este menu móvel conta com uma lista de ícones, cada um representando um tipo de mecânica a ser adicionada sendo, respectivamente, vídeo, áudio, texto, imagem, localização e composta. Este painel é, por padrão, desabilitado, estando transparente nesse estado. Para habilitá-lo, basta que uma única missão de qualquer seja selecionada. Com isso, será possível escolher o tipo a ser criado, ao clicar sobre algum, todos os outros painéis desaparecerão e espera-se que o mapa seja clicado no local que pretende-se adicionar o *marker* referente a nova mecânica como integrante da missão selecionada. Ao fazer isso e adicioná-la, ela estará disponível para fazer parte do jogo.

4.1.5 Botão de Adição. Este botão é voltado unicamente a função de adição dos novos elementos. Foi pensado de modo a proporcionar um local central onde o usuário tivesse o acesso rápido a esta funcionalidade, por mais que ela se encontre presente em outros locais.

4.1.6 Menu Lateral. O menu retrátil lateral contém opções gerais do jogo. Atualmente há: “configurações”, onde o game designer poderá optar pelo intervalo de salvamento automático, por exemplo; “publicação”, função a qual o jogo poderá ser publicado para

ser acessado pelo clientes móveis; “games”, página onde o usuário poderá ver a lista de jogos criados; e “Novo Jogo”, para a criação de um novo jogo desde o início.

5 CONCLUSÕES E TRABALHOS FUTUROS

Jogos Móveis Baseados em localização já se mostraram aplicações bastante atrativas ao longo dos últimos anos, com diversos títulos tendo alcançado um sucesso global. Estas aplicações tem potencial de alavancar outras áreas que não apenas a do entretenimento, como por exemplo, o turismo, esporte de aventura dentre outros.

A LAGARTO se insere neste contexto ao se apresentar como uma abordagem para que tais jogos possam se construídos por pessoas sem profundo conhecimento de desenvolvimento de aplicações e tecnologia. A partir de um estudo sobre as deficiências encontradas na primeira versão do editor de jogos, esta pesquisa de iniciação científica tem avançado em um projeto de refatoramento da interface gráfica deste editor. Os resultados apresentados já mostram uma interface mais limpa e concisa de ações possíveis para editores de jogos.

A conclusão desta pesquisa se dará com a finalização da ferramenta, e a realização de testes com usuários por meio de experimentos controlados para coleta da experiência destes usuários na nova versão do Editor.

REFERÊNCIAS

- [1] L.A. Lehmann. 2012. *Location-Based Mobile Games*. GRIN Verlag, Munich, Germany.
- [2] Carleandro Noletto. 2015. Uma ferramenta de autoria para o desenvolvimento de jogos móveis baseados em localização com realidade aumentada. (2015).
- [3] Janne Paavilainen, Hannu Korhonen, Kati Alha, Jaakko Stenros, Elina Koskinen, and Frans Mayra. 2017. The Pokémon GO Experience: A Location-Based Augmented Reality Mobile Game Goes Mainstream. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 2493–2498.
- [4] Luís Fernando Maia Silva, Carleandro Noletto, Messias Lima, Cristiane Ferreira, Cláudia Marinho, Windson Viana, and Fernando Trinta. 2017. LAGARTO: A Location based Games AuthoRing TOol enhanced with augmented reality features. *Entertainment Computing* 22 (2017), 3–13.
- [5] Heinrich Söbke, Jannicke Baalsrud Hauge, and Ioana A. Stefan. 2017. Prime Example Ingress Reframing the Pervasive Game Design Framework (PGDF). *International Journal of Serious Games* 4, 2 (June 2017), 39–58. <https://doi.org/10.17083/ijsg.v4i2.182>

Levantamento de tecnologias para ONGs de Proteção Animal para apoio ao resgate de animais domésticos acoplados ao ciclo de vida de um Sistema *Web*

Tatiana Tozzi
Instituto Federal Catarinense –
Campus Camboriú (IFC)
Rua Joaquim Garcia, s/n, Centro,
Camboriú, SC 88340-055
tatitozzi@hotmail.com

Daniel Fernando Anderle
Instituto Federal Catarinense –
Campus Camboriú (IFC)
Rua Joaquim Garcia, s/n, Centro,
Camboriú, SC - 88340-055
daniel.anderle@ifc.edu.br

Rodrigo Ramos Nogueira
Instituto Federal Catarinense –
Campus Camboriú (IFC)
Rua Joaquim Garcia, s/n, Centro,
Camboriú, SC - 88340-055
rodrigo.nogueira@ifc.edu.br

ABSTRACT

This paper reports on the stages of identification of existing technologies that can be used to identify and rescue lost animals and to disseminate animals for adoption by NGOs of Animal Protection, Zoonoses Center and Protective Guards. In this article the developed phases of the Course Completion Work are presented. For this, an exploratory research was carried out initially, followed by a questionnaire seeking to identify which technologies are used to identify, locate, rescue and adopt domestic animals. The current technologies found were Microchip - RFID, Microchip - NFC, Collar with qrCode, Collar with Tag, Search application, Identification application, Social networks. In this way the research reports the main technologies used to date and is going to develop an application proposal to improve and expand the dissemination of pets that are lost or for adoption.

KEYWORDS

Application model, identification of technologies, animal protection.

1 INTRODUÇÃO

Os animais domésticos fazem parte do dia-a-dia dos seres humanos desde os primórdios e são representados em maioria pelas espécies felina e canina. Cães e gatos são excelentes para serem companheiros dos seres humanos, além de proporcionarem conforto, amizade, e auxiliar no tratamento de doenças [4].

O Brasil possui a quarta maior população mundial de animais domésticos, segundo o IBGE de 2010, com 132,4 milhões de animais domésticos, sendo eles cães, gatos, aves, peixes e alguns animais exóticos [9]. Atualmente os animais domésticos são considerados como membros da família, conforme aponta o estudo realizado pela Associação Brasileira da Indústria de Produtos para Animais de Estimação (ABINPET) em 2016 [12].

Segundo a Organização Mundial da Saúde (OMS), conforme pesquisa realizada em 2014, estima que o Brasil possui 30 milhões de animais abandonados, sendo 10 milhões de gatos e 20 milhões de cachorros [2].

Os animais que se encontram em situação de rua provavelmente já fizeram parte de uma família e possuíam um lar, porém vieram a ser abandonados por seus próprios donos devido a questões socioeconômicas, religiosas ou culturais [9]. Devido a isso o abandono de cães e gatos tem aumentado a população de animais em situação de rua, passando a ser um desafio para o bem-estar dos animais e à saúde pública [8]. O uso da Internet e da tecnologia atualmente é a principal fonte de divulgação de eventos de adoção de animais e para localizar animais perdidos ou abandonados.

A partir dessa constatação, este trabalho vem sendo construído com o intuito de responder a seguinte pergunta de pesquisa: “Por meio de *software* é possível diminuir o abandono animal e melhorar a divulgação de animais para adoção na região da AMFRI¹”. Para atingir o objetivo este trabalho visa identificar as tecnologias utilizadas atualmente para auxiliar no resgate, identificação e divulgação de animais domésticos, tendo como base o seguinte roteiro:

- Apresentar os resultados através de uma Pesquisa de opinião;
- Descrever as tecnologias encontradas, classificá-las e identificar a aplicabilidade de uso;
- Desenvolver um projeto para o desenvolvimento de um Sistema *Web* para apoio às ONGs e Protetores Independentes;
- Testar a viabilidade do modelo proposto junto às ONGs e Protetores Independentes.

2 METODOLOGIA

Este trabalho se classifica quanto à natureza aplicada e tecnológica, em questão aos objetivos como exploratória e quanto aos procedimentos bibliográfica, conforme identificado

In: XV Workshop de Trabalhos de Iniciação Científica (WTIC 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pôsteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.
©2018 SBC – Sociedade Brasileira de Computação.
ISBN: 978-85-7669-435-9.

¹ Associação dos Municípios da Foz do Rio Itajaí – Santa Catarina, composta por 11 municípios (Balneário Camboriú, Balneário Piçarras, Bombinhas, Camboriú, Ilhota, Itajaí, Itapema, Luiz Alves, Navegantes, Penha e Porto Belo) [1].

por meio da Figura 1 – imagem A.

A metodologia adotada para o desenvolvimento deste trabalho foi dividida em seis fases, as quais estão relacionadas com os objetivos propostos desta pesquisa. A Figura 1 – imagem B, mostra as etapas que foram empregadas na metodologia de desenvolvimento deste trabalho.

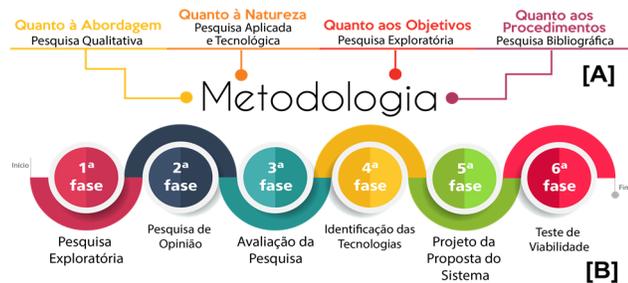


Figura 1: Metodologia [A] e Fases de desenvolvimento [B].

A primeira fase correspondeu na Pesquisa Exploratória, em bases de dados, livros, páginas *Web* e trabalhos relacionados buscando identificar tecnologias que possam auxiliar a localização de animais perdidos ou abandonados. Na segunda fase foi desenvolvido e aplicado a Pesquisa de Opinião na região da AMFRI. Na terceira foram avaliados os resultados da Pesquisa de Opinião. Na quarta fase consiste da identificação das tecnologias encontradas na Pesquisa de Opinião e nos Trabalhos relacionados.

A quinta fase será desenvolvido um projeto de Sistema *Web*, na qual serão estabelecidos os requisitos do sistema, os atores; desenvolvido os casos de uso, a modelagem do banco de dados e a prototipação das telas do sistema. A sexta fase constituirá de testar a viabilidade da aplicação proposta junto às ONGs, Centros de Zoonoses e Protetores Independentes, identificando se nossa proposta vem atender necessidades das mesmas para a divulgação, identificação, localização e adoção de animais domésticos.

3 FASES DE DESENVOLVIMENTO

Até o presente momento foram desenvolvidas quatro fases do trabalho, as quais serão apresentadas resumidamente a seguir.

3.1 Pesquisa exploratória

Esta fase teve como pretensão identificar os principais conceitos abordados durante o decorrer da pesquisa, os quais são: animais domésticos, posse responsável, Protetores Independentes, ONG, Centro de Zoonoses, resgate de animais, abandono de animais. Como parte desta fase foram selecionados seis trabalhos científicos (artigos e monografias). Por meio destes trabalhos buscou-se identificar outros projetos que foram desenvolvidos tendo como objetivo auxiliar os animais domésticos, desde um sistema informatizado para uma ONG, como a criação de um registro nacional de identificação de animais. Estes trabalhos e as tecnologias e a tecnologia proposta são apresentadas na subseção 3.4 – Identificação das tecnologias.

3.2 Pesquisa de opinião

Esta fase consistiu em desenvolver uma pesquisa de opinião com os moradores da região da AMFRI, por meio de um questionário composto por 24 perguntas utilizando o *Google Forms*. Antes da aplicação da pesquisa no público-alvo, uma busca foi realizada no *Google* e *Facebook* para localizar as ONGs e Protetores Independentes da região da AMFRI. Foram identificados 51 páginas e grupos no *Facebook* e 3 sites. Para cada um dos grupos, páginas e sites uma mensagem foi enviada com um convite para participar da pesquisa.

As perguntas tinham como objetivo identificar à faixa etária, o sexo, a quantidade de animais e a espécie de animais possuem, se os participantes da pesquisa fazem parte de alguma ONG de Proteção Animal ou se atua como Protetor Independente, quais tecnologias já usaram para divulgar animais abandonados ou para adoção entre outras questões.

3.3 Avaliação da pesquisa

Esta fase apresenta os principais resultados da pesquisa. Após duas semanas de divulgação e aplicação da pesquisa de opinião, 100 pessoas responderam à pesquisa. As cidades que mais participaram da pesquisa foram as cidades, Camboriú com 45% e Balneário Camboriú com 32%. A maioria dos participantes tem entre 26 a 31 anos (24%), sendo que 69% das participações foram de mulheres. 89% afirmaram ter animais domésticos.

Dos participantes, 71% já utilizaram alguma tecnologia para auxiliar no resgate e divulgação de animais domésticos. A pergunta em seguida continha várias tecnologias², a qual os participantes poderiam selecionar quais já usaram, sendo a mais identificada as Redes Sociais. Questionados se gostariam de conhecer as tecnologias citadas na pesquisa, 93% responderam que sim, uma vez que 61% dos participantes não conheciam as tecnologias citadas. Por meio da 19ª pergunta, descobrimos que 77% dos participantes utilizam redes sociais para auxiliar na divulgação de animais perdidos, abandonados ou para adoção. Sendo o *Facebook*, a rede social mais utilizada para este fim e as publicações desses anúncios são em maioria feitas na própria linha do tempo dos participantes. 98% dos participantes acham que um Sistema *Web* poderia melhorar e ampliar a divulgação de animais perdidos ou para adoção.

Como forma de devolutiva para os participantes da pesquisa foram elaborados dois infográficos, o primeiro contendo as respostas utilizando gráficos e o segundo (figura 2) apresenta os principais dados das respostas. Ambos infográficos foram enviados por e-mail aos participantes e publicado no *Facebook* como forma de agradecimento.

3.4 Identificação das Tecnologias

Nesta fase foram abordadas as tecnologias identificadas na Pesquisa de opinião e nos Trabalhos relacionados.

3.4.1 Tecnologias identificadas na Pesquisa de Opinião.

² *Microchip* - RFID, *Microchip* - NFC, Coleira com *qrCode*, Coleira com *Tag*, Aplicativo de busca, Aplicativo de Identificação, Redes sociais.

O **microchip RFID** ou *Radio-Frequency IDentification* (Identificação por rádio frequência) é um método de identificação automática por meio de sinais de rádio. Os sinais são recuperados e armazenados dados remotamente por meio de um dispositivo de *tags* ou etiqueta RFID. A *tag* ou etiqueta é um *transponder* ou microchip que pode ser colocado em uma pessoa, animal, produto ou embalagem. **Microchip NFC** (*Near Field Communication*) é uma tecnologia que possibilita a troca de informações e dados entre dispositivos. Para que ocorra a troca de informações é apenas necessário que ambos dispositivos estejam até 10 centímetros próximos um do outro, tornando essa transmissão de dados segura [5].

colocado na coleira do animal, e caso ele venha se perder pode-se utilizar um leitor para obter essa informação e promover o reencontro entre animal e tutor.

Aplicativos de busca podem ser utilizados para cadastrar animais que tenham fugido de casa ou se perderam de seus tutores. Através deste cadastro os tutores inserem fotos e informações do animal desaparecido e dados de contato no aplicativo. Já **Aplicativo de identificação** promovem a identificação do animal por meio de reconhecimento facial, utilizando a tecnologia de comparação de imagens (visão computacional e inteligência artificial). Como citado anteriormente a **rede social** mais utilizadas pelos brasileiros é o *Facebook* e além desta rede social as demais redes também são grandes aliadas na procura de animais perdidos, porém em fevereiro de 2018 foi criada da rede social *Puppyfi*, com o principal objetivo auxiliar os animais, assim auxiliando tutores a encontrarem seus animais desaparecidos e encaminhar os que foram abandonados para adoção.

3.4.2 Tecnologias identificada nos trabalhos relacionados.

No artigo de Carpanezi, Tomazela e Pontes [3] apresentam o desenvolvimento de um aplicativo, onde foram cadastrados os animais para adoção. A aplicação possibilita quem tivesse interesse em adotar algum dos animais cadastrados possa conhecer as características do mesmo de forma fácil e rápida.

Já Silva, Franco e Molina [13] teve como proposta à implementação de um sistema informatizado para registro de controle dos animais acolhidos pelo Centro de Controle de Zoonoses. O Projeto Adoção Animal de Evangelista et al. [7] teve como pretensão ampliar a divulgação dos animais abandonados que vivem em abrigos utilizando mídias sociais [*Facebook* e *Blogger*] e outros recursos físicos, buscando incentivar a adoção de animais.

No trabalho de Menezes Filho e Souza [11] teve como objetivo o desenvolver uma ferramenta para registro e identificação de animais de companhia. Foram utilizadas as tecnologias para o desenvolvimento web (*JHipster*, *Spring*, *Angular* e *Bootstrap*), como a intenção de criar uma base de dados para as prefeituras armazenarem informações sobre a população de animais e promover ações estratégicas para controle da populacional.

Na monografia de Donatti [6] foram desenvolvidos dois sistemas utilizando um protocolo baseado no *ZigBee* e a tecnologia *Global Positioning System* – GPS. Ambos sistemas foram criados para coletar dados referentes ao posicionamento global dos animais (bovinos), identificando a longitude e latitude do animal.

3.4.3 Classificação das tecnologias. Conforme a tabela 1, a coluna 1 contém as tecnologias encontradas, na coluna 2 e 3 foram assinalados utilizando os ícones  para não e  para sim, as tecnologias classificadas como *Hardware* e *Software* respectivamente. Algumas das tecnologias listadas na tabela 2, são consideradas *hardware* e *software* pois uma demanda da outra para total funcionalidade da tecnologia.



Figura 2: Infográfico com os dados principais da pesquisa.

A **coleira com qrCode** (*Quick Response Code*), consiste de uma coleira com uma medalha de identificação com *qrCode*. A medalha de alumínio possui um *qrCode* que possibilita acessar a página do animal. Por meio da página a pessoa que venha encontrar o animal com esta coleira pode ter acesso as informações de contato do tutor, telefone do médico veterinário, fotos e informações médicas. Além do *qrCode* a medalha pode conter o endereço (*link*) da página do animal, caso quem tenha encontrado o animal não tenha acesso a um dispositivo que lê o *qrCode*. A **coleira com tag** é uma alternativa para utilização do *microchip* sem que este seja implantado no animal. A *tag* contendo os dados do animal (o código de identificação) é

Tabela 1. Classificação das tecnologias

Tecnologias	Hardware	Software
Microchip RFID	✓	✓
Microchip NFC	✓	✓
Coleira com qrCode	✗	✓
Coleira com Tag	✓	✓
Aplicativo de Busca	✗	✓
Aplicativo de Identificação	✗	✓
Redes sociais	✗	✓
Aplicativo mobile para adoção	✗	✓
Informatização de Centro de Zoonoses	✗	✓
Facebook e Blogger como ferramenta de divulgação	✗	✓
Registro geral de animais	✗	✓
Monitoramento de animais com rede de sensores sem fio	✓	✓

3.4.4 Aplicabilidade das tecnologias

Segundo a tabela 2, a coluna 1 representa as tecnologias encontradas, na coluna 2 as ONGs, na coluna 3 os Centros de Zoonoses e na coluna 4 os Protetores Independentes. Foram assinados utilizando os ícones ✗ para não e ✓ para sim.

Tabela 2. Aplicação das tecnologias

Tecnologias	ONGs	Centros de Zoonoses	Protetores Independentes
Microchip RFID	✓	✓	✓
Microchip NFC	✓	✓	✓
Coleira com qrCode	✗	✗	✓
Coleira com Tag	✗	✗	✓
Aplicativo de Busca	✓	✓	✓
Aplicativo de Identificação	✓	✓	✓
Redes sociais	✓	✓	✓
Aplicativo mobile para adoção	✓	✓	✓
Informatização de Centro de Zoonoses	✓	✓	✗
Facebook e Blogger como ferramenta de divulgação	✓	✓	✓
Registro geral de animais	✓	✓	✓
Monitoramento de animais com rede de sensores sem fio	✓	✓	✗

4 CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

Espera-se que este trabalho possa contribuir para a conscientização da importância da adoção responsável de animais domésticos e no desenvolvimento de uma aplicação que possa auxiliar o resgate, localização, identificação e adoção de animais domésticos.

Este artigo visa apresentar as etapas que foram realizadas e as que serão desfeitas no decorrer do Trabalho de Conclusão de Curso. Assim, o *feedback* da comunidade do *Workshop* de Trabalhos de Iniciação Científica é de grande importância para refinar e aprimorar esta pesquisa, assim levando a uma conclusão satisfatória. Já como trabalhos futuros pretende-se dar seguimento as duas fases finais (5 e 6), e iniciar o desenvolvimento da aplicação.

REFERÊNCIAS

- [1] AMFRI Associação dos Municípios da Foz do Rio Itajaí. Acessado em 7 Mai. 2018 de <https://www.amfri.org.br/>
- [2] ANDA. Brasil tem 30 milhões de animais abandonados. 2014. Acessado em 14 Set. 2017 de <https://anda.jusbrasil.com.br/noticias/100681698/brasil-tem-30-milhoes-de-animais-abandonados>
- [3] CARPANEZI, C. A., et al. Desenvolvimento de um aplicativo mobile para adoção de animais de estimação. 2016. Acessado em 28 fev. 2018 de <http://reverte.fatecid.com.br/index.php/revista/article/view/183>
- [4] CARVALHO, L. 9 benefícios que bichos de estimação trazem à saúde. 2016. Acessado em 13 Set. 2017 de <https://exame.abril.com.br/estilo-de-vida/9-beneficios-que-bichos-de-estimacao-trazem-a-saude/>
- [5] COIMBRA, D. da S. O Uso da Tecnologia NFC na Identificação PET. 2016. Acessado em 28 fev. 2018 de <http://www2.uesb.br/computacao/wp-content/uploads/2014/09/O-Uso-da-Tecnologia-NFC-da-Identificacao-C3%A7%C3%A3o-PET-DIEGO-COIMBRA.pdf>
- [6] DONATTI, R. N. Desenvolvimento de um sistema de monitoramento de animais, utilizando rede de sensores sem fio, baseado no protocolo ZigBee e tecnologia GPS. 2017. Acessado em 10 mai. 2018 de <http://www.teses.usp.br/teses/disponiveis/74/74134/tde-14062017-112910/publico/ME8845716COR.pdf>
- [7] EVANGELISTA, A. G., et al. Projeto adoção animal: Incentivando a prática da adoção de Cães e Gatos Abandonados – Resultados preliminares. 2015. Acessado em 10 mai. 2018 de <http://eventos.ifc.edu.br/wp-content/uploads/sites/5/2015/10/PROJETO-ADO-C3%87%C3%83O-ANIMAL-IFC-INCENTIVANDO-A-PR-C3%81TICA-DA-ADO-C3%87%C3%83O-DE-C3%83ES-E-GATOS-ABANDONADOS-RESULTADOS-PRELIMINARES.pdf>
- [8] GARCIA, R. C. M. Normas e políticas públicas para controle populacional de cães e gatos. In: CONGRESSO BRASILEIRO DE BIOÉTICA E BEM-ESTAR ANIMAL, 3., 2014, Curitiba: Anais...Curitiba: UFPR/LABEA, 2014. p. 149.
- [9] IBGE. Um panorama da saúde no Brasil: acesso e utilização dos serviços, condições de saúde e fatores de risco e proteção à saúde. 2010. Acessado em 13 nov. 2017 de http://bvsms.saude.gov.br/bvs/publicacoes/pnad_panorama_saude_brasil.pdf
- [10] MELO, L. Como o brasileiro cuida e quanto gasta com seus pets. 2016. Acessado em 21 Out. 2017 de <https://exame.abril.com.br/negocios/como-o-brasileiro-cuida-e-quanto-gasta-com-seus-pets/>
- [11] MENEZES FILHO, C. T. B et al. Registro geral de Animais (RGA): um sistema para o registro e identificação de animais de companhia. 2017. Acessado em 12 abr. 2018 de <https://repositorio.ufsc.br/xmlui/bitstream/handle/123456789/176967/RGA.pdf>
- [12] REZENDE, L. F. G., et al. Perfil dos proprietários de cães e gatos e a prática da guarda responsável dos acadêmicos CEULJI-ULBRA. Archives of Veterinary Science, v. 17, p.34-36, resumo 012, 2012
- [13] SILVA, et al. Sistema Informatizado para o Centro de Controle de Zoonoses – Americana/SP. 2017. Acessado em 10 mar. 2018 de https://febrace.org.br/mostra3m2017/_data/pdf/3M_anaisVMostra.pdf/

WebMedia 2018

XXIV Simpósio Brasileiro de Sistemas Multimídia e Web

ANAIS

**XVII Workshop de Ferramentas
e Aplicações**

Coordenadores

Carlos de Salles (UFMA)

Álan L.V. Guedes (PUC-Rio)

XVII Workshop de Ferramentas e Aplicações

Sumário

Prefácio	89
Organização	90
Sessão Técnica 1. Sistemas Web e Internet de Coisas	91
Towards Developing a Practical Tool to Assist UX Evaluation in the IoT Scenario	91
Rodrigo L.A. Almeida (UFC), Ticianne G.R. Darin (UFC), Rossana M.C. Andrade (UFC), Italo L. de Araújo (UFC)	
OpenData Processor: An Automation tool for the process of extracting and publishing open data to CKAN	97
Allyson Vilela (IFRN), André Almeida (IFRN), Frederico Lopes (UFRN)	
MobSink: a Visual Mobile Wireless Sensor Networks Positioning Simulator	103
João Paulo Just Peixoto (IFBA), Daniel G. Costa (UEFS)	
Integração de um Gerador de Código ao FrameWeb Editor	109
Breno Leite Zupeli (UFES), Vítor E. Silva Souza (UFES)	
Sessão Técnica 2. Video, Realidade Aumentada e Machine Learning	115
VolleyJump: Uma aplicação para a análise de saltos no voleibol de praia	115
Renan Bandeira (UFC), Fernando Trinta (UFC), João Gomes (UFC), Marcio Maia (UFC), Alexandre Araripe (UFC)	
QuiRA: An Augmented Reality Application to Support Chemistry Learning	121
Letícia Porto Soares (IFBA), Camilo Alves Carvalho (IFBA), Luis Paulo da Silva Carvalho (IFBA), Manuela Amaral de Araújo (IFBA), Paulo Henrique Teixeira Silva (IFBA), Wdson Costa (IFBA)	
MCU em software distribuído em nuvem	127
Valter Roesler (UFRGS), Mário Gasparoni Júnior (UFRGS), Ronaldo Husemann (UFRGS), Roberto Irajá Tavares da Costa Filho (UFRGS), Rafael Valle (RNP)	
Sistema Emagreça@Saudável para controle da obesidade em adultos	133
Maria Lúcia Kroeff Barbosa (UFRGS), Luciano Zancan (UFRGS), Mario Figueiró Zemor (UFRGS), Valter Roesler (UFRGS), Marcia Rosa da Costa (UFCSPA), Sílvio César Cazella (UFRGS/UFCSPA)	

Sessão Técnica 3. Dispositivos Móveis e Jogos	139
Digital Cognitive Tests with Brazilian elderly	139
Leandro Agostini do Amaral (USP), Renata Pontin de Mattos Fortes (USP), Thiago Jabur Bittar (UFG)	
Authoring of interventions with on-demand generation of screen for mobile devices	145
Bruna C.R. Cunha (USP/IFSP), Kamila R.H. Rodrigues (USP), Leonardo F. Scalco (USP), Isabela Zaine (USP), Caio C. Viel (USP), Maria da Graça C. Pimentel (USP)	
Ballgorithm - Uma Ferramenta Introdutória Para Conceitos de Programação	151
Guilherme Afonso Melo Sousa Melo (UFMA), Carlos de Salles Soares Neto (UFMA)	

XVII Workshop de Ferramentas e Aplicações

Prefácio

O XVII Workshop de Ferramentas e Aplicações (WFA) é um dos eventos integrantes do Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia) e constitui um importante fórum dedicado à apresentação e discussão de ferramentas e aplicações desenvolvidas nas áreas de Multimídia, Hipermídia e Web. O evento visa integrar pessoas com interesses comuns, criar redes de pesquisa e desenvolvimento, e dar visibilidade aos potenciais pesquisadores junto às comunidades acadêmica e industrial.

Neste ano, o WFA recebeu um total (recorde) de 22 submissões de ferramentas e aplicações, de instituições do Brasil inteiro e de áreas diversas. Os artigos foram avaliados de forma criteriosa por um Comitê de Programa (CP) de alto nível, com pelo menos 3 revisões cada. Uma inovação nesta edição foi a análise, pelos membros do CP, do potencial de negócios das ferramentas e aplicações submetidas. Ao final do processo de revisão, foram selecionadas 11 ferramentas ou aplicações para apresentação nas sessões do WFA, demonstração ao público em momentos específicos durante o coffee-break e publicação nos anais do evento.

Agradecimentos a todos os autores que submeteram artigos para o WFA, contribuindo para o seu enriquecimento. Agradeço também a todos os membros do comitê de programa por sua dedicação, pelas valiosas revisões dos artigos e pelo comprometimento com os prazos estabelecidos.

Agradecimento especial ao Professor Manoel C.M. Neto (IFBA) pelo convite para coordenar o WFA e pela cooperação durante todo o processo e ao Professor Roberto Willrich (UFSC) pela coordenação de publicações do WebMedia.

Salvador, outubro de 2018.

Carlos de Salles (UFMA)

Álan L.V. Guedes (PUC-Rio)

Coordenadores do XVII Workshop de Ferramentas e Aplicações

Coordenador do XVI WFA

Carlos de Salles (UFMA)

Álan L.V. Guedes (Telemídia/PUC-Rio)

Comitê de Programa do XVI WFA

Adriano César Pereira (UFMG)

André Damasceno (PUC-Rio)

Antonio de Carvalho Jr (USP)

Antonio Busson (PUC-Rio)

Carlos Ferraz (UFPE)

Daniel Souza (UFERSA)

Débora Muchaluat-Saade (UFF)

Eduardo Albuquerque (UFG)

Ewerton Salvador (UFPB)

Fabio de Jesus Lima Gomes (IFPI)

Fernando Trinta (UFC)

Frank Siqueira (UFSC)

Frederico Durão (UFBA)

Glauco Amorim (CEFET-RJ)

Graça Bressan (USP)

Guilherme Lima (PUC-Rio)

Joberto Martins (UNIFACS)

Joel dos Santos (CEFET-RJ)

Lívia Ruback (UFRJ)

Leandro Souza (UFERSA)

Lincoln David Nery e Silva (UFPB)

Marcos Roriz Junior (UFG)

Marcos Seruffo (UFPA)

Mario Teixeira (UFMA)

Mario Godoy Neto (UNIVASF)

Marx Viana (PUC-Rio)

Renata Galante (UFRGS)

Ricardo Mendes (UVV)

Roberto Gerson Azevedo (EPFL)

Rodrigo Laiola Guimarães (UFES)

Ruan Gomes (IFPB)

Rudinei Goularte (USP)

Tatiana Tavares (UFPE)

Thaís Gaudêncio do Rêgo (UFPB)

Valdecir Becker (UFPB)

Windson Viana (UFC)

Yuri Malheiros (UFPB)

Towards Developing a Practical Tool to Assist UX Evaluation in the IoT Scenario

Rodrigo L. A. Almeida
Universidade Federal do
Ceará
Fortaleza, Brasil
rodrigoalmeida@great.ufc.br

Ticianne G. R. Darin
Universidade Federal do
Ceará
Fortaleza, Brasil
ticianne@virtual.ufc.br

Rossana M. C. Andrade
Universidade Federal do
Ceará
Fortaleza, Brasil
rossana@ufc.br

Italo L. de Araújo
Universidade Federal do
Ceará
Fortaleza, Brasil
italoaraujo@great.ufc.br

ABSTRACT

Internet of Things (IoT) is a paradigm that offers the users a novel and intelligent way of interacting with everyday objects. However, the complexity of an IoT scenario, involving the interaction among multiples devices, users and services, represents a challenge for the evaluation of user experience (UX). Hence, to assist in the conduction of a more effective UX evaluation in an IoT environment, this paper presents a checklist that assists evaluators with different levels of experience to perform UX evaluation in the IoT scenario. Such tool was built based on three main steps: a literature review, the UX evaluation of an IoT application, and a questionnaire answered by IoT experts.

KEYWORDS

Internet of Things, User Experience Evaluation and User Observation.

1 INTRODUÇÃO

A evolução tecnológica de áreas do conhecimento como redes de sensores sem fio, computação em nuvem e computação ubíqua permitiu o surgimento da Internet das Coisas, do inglês *Internet of Things* (IoT) [14]. IoT permite a presença pervasiva de uma variedade de coisas ao redor das pessoas. Esses objetos possuem um esquema de endereçamento único que permite a interação com outros objetos para alcançar objetivos comuns [2]. Assim, atividades do cotidiano como cozinhar, acender luzes, fazer compras podem ser auxiliadas por objetos embutidos de poder computacional e comunicação [9].

Dessa forma, em um ambiente IoT, o usuário é capaz de interagir com diversas “coisas” simultaneamente, o que resulta em uma interação complexa para o usuário. Isso ocorre porque ações e informações necessárias para as aplicações no ambiente IoT não são, muitas vezes, claras e facilmente compreendidas e utilizadas pelos usuários [11]. Por esse motivo, a literatura tem apontado que a avaliação da experiência do usuário, do inglês

User Experience (UX), no cenário de IoT traz desafios relacionados à natureza multi-plataforma da interação, bem como à interconexão entre dispositivos, serviços e usuários [5,12]. A avaliação de UX no paradigma IoT é complexa porque precisa considerar aspectos peculiares a esses ambientes, como o número significativo de “coisas” presentes na avaliação.

Dentre esses aspectos, Roland et al. (2015) [11] destacam, por exemplo, que a interação com uma funcionalidade pode estar distribuída entre múltiplos dispositivos, o que inclui várias interfaces, usuários e regras. Além disso, o foco da UX pode ser no serviço ao invés da interação com um dispositivo, incluindo diversos tipos de contexto de uso. Além disso, tanto IoT, como UX apresentam aspectos, que no contexto da avaliação de um ambiente inteligente precisam estar estruturados para a identificação da experiência do usuário [5].

Com o intuito de auxiliar a comunidade científica a conduzir a avaliação de UX no cenário IoT, este trabalho propõe uma ferramenta de avaliação no formato de um *checklist* para apoiar a observação do comportamento do usuário durante a interação com uma aplicação IoT. O *checklist* foi construído com base na metodologia de Stufflebeam (2000) [13] com o apoio dos resultados de uma revisão da literatura, de uma avaliação de UX em um cenário IoT e de um questionário com especialistas de IoT. A ferramenta indica a possibilidade de auxiliar pesquisadores com pouca experiência na avaliação de UX no cenário de IoT.

2 FUNDAMENTAÇÃO TEÓRICA

A IoT é um paradigma que permite que pessoas e objetos inteligentes estejam conectados a qualquer momento, lugar, seja qual for o objeto, contanto que estejam utilizando uma rede e seus serviços [15]. A IoT tem apresentado forte influência no desenvolvimento tecnológico de vários domínios como cidades inteligentes, casas e escritórios inteligentes e na área da saúde [6]. Através desses domínios, é possível perceber a diversidade de aplicações com as quais o usuário pode interagir, resultando em tipos de experiências distintas em cada um desses sistemas.

Dessa forma, a interação em sistemas IoT apresenta desafios devido à forma como esta ocorre no cenário de IoT, bem como no contato entre o usuário e o próprio sistema. A interação é um dos desafios que surge com a IoT [1,3] e ela difere na forma como a ocorre no sistema e na forma do contato do usuário com o sistema IoT. Andrade et al. (2017) [1] e Rowland et al. (2015) [11] abordam a interação no contexto da IoT em dois níveis:

In: XVII Workshop de Ferramentas e Aplicações (WEA 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pôsteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.
©2018 SBC – Sociedade Brasileira de Computação.
ISBN: 978-85-7669-435-9.

Thing-Thing Interaction e *Human-Thing Interaction*. O primeiro se relaciona à intercomunicação entre os objetos inteligentes com o intuito de trocar informações e fornecer serviços aos usuários. O segundo nível se refere à interação que ocorre entre o objeto inteligente e o usuário.

A interação da *thing-thing* pode ser organizada em dois níveis: o de Internet e o de coisas. As questões relacionadas à Internet são centradas na conexão e relacionadas a aspectos como sincronia dos serviços, responsividade e confiabilidade da rede. Quanto ao nível das coisas, são abordados aspectos como consumo de energia, sensibilidade ao contexto, interoperabilidade e dificuldade de instalação. Todas as características dos objetos inteligentes da interação *thing-thing* impactam diretamente na interação do usuário com o sistema IoT [1], que consequentemente tem influência na UX, pois afetam a forma como o usuário percebe o sistema e como são atendidas suas expectativas e necessidades.

Andrade et al. (2017) [1] e Bergman e Johansson (2017) [3] listam ainda desafios na avaliação da interação na perspectiva do usuário. Segundo os autores, ambientes e sistemas IoT podem gerar diversos cenários de uso devido às possíveis combinações de objetos inteligentes, serviços e interfaces.

3 PROCESSO DE DESENVOLVIMENTO DO CHECKLIST

O *checklist* de observação do comportamento do usuário foi desenvolvido baseado nos passos da metodologia proposta por Stufflebeam (2000) [13], a qual organiza em doze fases a elaboração de um *checklist*. As três primeiras fases se concentram na definição do domínio, geração dos itens de verificação, a quarta e quinta fases abordam a geração das categorias e da fase seis a fase doze são realizadas a revisão, o refinamento, a formatação e a validação do *checklist*.

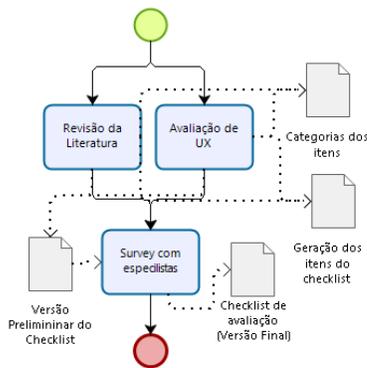


Figura 1: Passos para a construção do *checklist*.

Neste trabalho, aplicamos os princípios gerais da metodologia de Stufflebeam (2000) [13] que consistem na geração dos itens, classificação em categorias e a participação de especialistas de domínio na revisão e validação do instrumento. Para a geração e organização dos itens de verificação utilizamos os resultados de uma revisão da literatura e uma avaliação de UX

em um ambiente IoT nas quais foi possível observar aspectos práticos relacionados com a condução da avaliação. Com a triangulação dos resultados desses métodos, aplicamos um questionário com especialistas de IoT que auxiliaram na revisão e validação do *checklist*. A Figura 1 descreve os métodos aplicados na geração do instrumento de avaliação e os principais resultados de cada um.

3.1 Revisão da Literatura

O objetivo da revisão literatura foi a identificação de estudos que realizassem a avaliação UX no contexto de IoT, para identificação de métodos de avaliação e características de IoT relevantes no contexto de UX. Para isso, foi utilizada a string descrita na Figura 2. A *string* foi aplicada na *Scopus Digital Library* resultando em um conjunto de 471 estudos. Todo o processo da revisão foi executado por um especialista em avaliações de IHC (Interação Humano-Computador), com cinco anos de experiência na área, e revisado por um especialista de IoT e outro de IHC. Após a leitura de título e resumo, seguido pela leitura completa dos trabalhos foram selecionados um conjunto de 17 estudos. Devido ao limite de espaço do artigo, os estudos identificados na revisão estão disponíveis em: <https://goo.gl/KBSTQ9>.

```
TITLE-ABS-KEY ((evaluation OR assessment) AND
("user experience" OR usability)) AND
("Internet of thing*" OR "smart home*" OR
"smart building*" OR "smart environment*").
```

Figura 2. *String* de busca utilizada na revisão da literatura.

Os estudos apresentaram a predominância de utilização de métodos empíricos (94%), como observação do comportamento do usuário e entrevista, e a avaliação de características de IoT como *programabilidade* (65%) e *sensibilidade ao contexto* (29%). Foram identificados alguns problemas de UX recorrentes em aplicações IoT, assim como as principais dificuldades descritas pelos avaliadores. Na Tabela 1 observamos a descrição de alguns exemplos encontrados na literatura de problemas e dificuldade no cenário descrito anteriormente.

Tabela 1: Exemplo de descrição dos problemas e dificuldades identificado na revisão da literatura

Problemas de UX		Dificuldades na execução da avaliação	
Estudo	Descrição	Estudo	Descrição
[10]	Dificuldades na instalação do objeto inteligente no ambiente IoT.	[4]	Avaliação no ambiente controlado pode destoar do ambiente de uso real do usuário.
[8]	Dificuldade na identificação das funcionalidades distribuídas no ambiente e no mapeamento delas na interface do dispositivo móvel.	[4]	A UX pode ser diferente de um sistema IoT avaliado individualmente, em relação a um ambiente com várias aplicações IoT.

A partir dos problemas de UX no cenário de IoT identificados na revisão da literatura, iniciamos a elaboração dos itens de

verificação de um *checklist* de observação do comportamento do usuário. O *checklist* de observação foi escolhido por ser uma ferramenta que pode ser empregada com métodos empíricos e auxiliar os avaliadores na condução da avaliação.

Dessa forma, a Tabela 1 exemplifica os tipos de problemas e dificuldades listados na literatura, que deram insumo para a primeira versão do instrumento de avaliação. O foco deste estudo não é ser exaustivo na discussão dos resultados da revisão da literatura, mas na descrição dos aspectos que auxiliaram na construção do instrumento de avaliação de UX no cenário de IoT.

3.2 Avaliação da UX em um ambiente IoT

Em paralelo à revisão da literatura, foi realizado uma avaliação de UX com uma aplicação IoT. O principal objetivo desta etapa foi verificar as dificuldades ao realizar a avaliação de UX em um ambiente IoT e quais problemas de UX ocorreriam na aplicação IoT sob avaliação. Isso auxiliou principalmente na visualização das categorias e melhor organização dos aspectos envolvendo UX e IoT. Para alcançar isso, foi utilizado o Automa GREat [1], um aplicativo móvel que realiza o controle de luzes inteligentes (três luzes) e um ar condicionado do ambiente foi avaliado. A Figura 3 apresenta as interfaces do Automa GREat.



Figura 3: Interfaces do aplicativo Automa GREat.

A avaliação foi organizada em três partes *i) assinatura do termo de consentimento e coleta de dados demográficos e experiência com tecnologia por meio de questionário, ii) observação do comportamento do usuário ao realizar tarefas específicas e iii) entrevistas e questionário final.*

Após a realização de um teste piloto, a avaliação foi realizada por cinco usuários, sendo quatro homens e uma mulher que executaram um conjunto de seis tarefas em laboratório. Os usuários tinham de 18 a 60 anos, com familiaridade com dispositivos móveis e precisavam saber mexer nas funções básicas de um ar-condicionado.

Os resultados da avaliação são apresentados na Tabela 2. A partir deles, foi possível observar quais aspectos relacionados a atenção do observador deve estar concentrada. Diferente de aplicativos tradicionais, a interface transpõe um único dispositivo e se torna o ambiente como um todo. Tal cenário pode vir à torna a observação da interação mais custosa devido à quantidade de elementos a serem observados. Além disso, foi possível observar como a interação *thing-thing* pode impactar na UX. Por exemplo, ao apresentar falha na conexão com o ar-condicionado, 40% dos usuários não conseguiram concluir a T5.

Além disso, o aplicativo não emitia nenhum *feedback* acerca do problema ao usuário tornando um problema crítico.

Tabela 2: Tarefas realizadas pelos usuários, taxa de completude e exemplos de problemas identificados.

Tarefas	% Taxa de Completude	Exemplo de problema de UX e IoT identificado
T1. Ligar ar-condicionado.	80%	O usuário realizou os passos da tarefa corretamente. No entanto, a aplicação falhou em ligar o ar condicionado causando frustração ao usuário.
T2. Ligar luzes.	100%	O usuário ficava inseguro quanto ao feedback de qual das luzes seria acesa no ambiente ao interagir na interface do aplicativo.
T3. Mudar a temperatura do ar-condicionado.	80%	O usuário realizou os passos da tarefa corretamente. No entanto, a aplicação falhou em ligar o ar condicionado causando frustração ao usuário.
T4. Alterar a cor das luzes.	80%	O usuário não conseguiu identificar como alterar a cor das lâmpadas.
T5. Desligar ar-condicionado.	60%	O usuário realizou os passos da tarefa corretamente, mas a aplicação não funcionou ao comando do usuário.
T6. Desligar luzes.	100%	O usuário teve dificuldade no mapeamento da interação de qual luz seria apagada no ambiente ao interagir com a interface.

3.3 Questionário com Especialistas de IoT

Com base nos resultados da revisão da literatura e da avaliação da UX em um ambiente IoT, um conjunto de 15 itens de verificação para avaliação da UX em ambientes IoT foi proposto. Para refinar e avaliar tais itens, foi enviado um questionário para profissionais e pesquisadores que desenvolvem e avaliam aplicações no contexto IoT. O questionário foi disponibilizado online e era constituído por 14 questões (quatro abertas) organizadas em três seções. A primeira seção abordava questões sobre dados demográficos do especialista; a segunda, sua experiência com IoT; e a terceira, investigava sua percepção sobre um conjunto de itens de verificação gerados com resultados dos métodos apresentados anteriormente neste trabalho. É importante reforçar que aspectos éticos de privacidade e anonimato foram respeitados durante a aplicação do questionário.

Dos 14 respondentes, 64% apresentava de um a dois anos trabalhando com IoT e 36% apresentavam experiência de quatro a cinco anos na área. Ainda relacionado à experiência com IoT, os participantes apresentaram uma média de participação do desenvolvimento e/ou avaliação de pelo menos duas aplicações no cenário de IoT, em domínios como *smarthomes, mobilidade urbana* e saúde. Os especialistas selecionaram os itens de verificação mais importantes ao realizar uma avaliação de UX em um ambiente IoT. Na Tabela 3 são apresentados os pontos do *checklist* melhor avaliados pelos especialistas. Além da seleção, os especialistas sugeriram melhorias na descrição dos itens e a incorporação de aspectos não contemplados pelos pontos de verificação apresentados no questionário. Dessa forma, os itens

de verificação passaram a incorporar melhor e de maneira mais clara aspectos relacionados à *programabilidade* e *sincronia* de serviços, estratégias de interação, satisfação e prazer no uso.

Tabela 3: Avaliação dos itens de verificação pelos especialistas de IoT.

%especialistas	Item de verificação (<i>sensibilidade ao contexto / human-thing</i>)
80%	O sistema IoT percebe as mudanças de contexto e se adequa corretamente as necessidades do usuário.
70%	O usuário percebeu problemas de funcionamento relacionados a interação <i>thing-thing</i> (Por exemplo: Problemas de sincronia, responsividade e conectividade das “coisas”).
60%	O usuário demonstra sinais de desconforto ou incômodo com o sistema IoT (Sinais de cansaço, insegurança durante a interação).
60%	O usuário demonstra sinais de frustração durante a interação (Por exemplo: expressões de descontentamento ou tristeza).

3.4 Discussão

Todas as situações apresentadas anteriormente pela aplicação de cada método auxiliaram na construção e definição do instrumento de avaliação. A revisão da literatura e avaliação de UX auxiliaram no levantamento dos primeiros pontos de verificação relacionados principalmente as interações (*thing-thing* e *human-thing*) e as características de IoT que vem ganhando destaque como, *sensibilidade ao contexto* e *programabilidade*. Além disso, a definição por um *checklist* seguiu a necessidade de observar o comportamento do usuário como principal meio de obter evidências sobre a UX em um cenário de IoT. Tais percepções capturadas pela literatura foram também identificadas na avaliação do Automa GREat.

O questionário com especialistas em IoT auxiliou na organização, na priorização e melhoria dos itens do *checklist*. É importante observar que o conjunto final itens de verificação são mais voltados ao contexto de *casas e prédios inteligentes*. Isso ocorreu devido a forte influência dos estudos da revisão da literatura, da avaliação de UX e do conhecimento maior dos especialistas em soluções nesse domínio. Tal cenário não limita a avaliações de soluções nesse domínio, mas evidência a fase em desenvolvimento em que se encontra a ferramenta. Dessa forma, é necessário que o instrumento seja aplicados em avaliações de UX no cenário IoT para que possa se obter as melhorias necessárias ao *checklist*.

4 CHECKLIST PARA AVALIAÇÃO DE UX NO CENÁRIO DE IOT

A ferramenta apresenta 19 itens de verificação organizados em duas categorias de interação (*human-thing* e *thing-thing*) presentes em um ambiente IoT, como discutido por Andrade et al. (2017) [1] e Roland et al. (2015) [11]. Essa organização foi escolhida por auxiliar os avaliadores na imersão e na compreensão do ambiente IoT pela perspectiva do usuário. Além disso, os problemas e melhorias identificados por cada categoria podem facilitar a evolução do sistema IoT por parte dos desenvolvedores.

A categoria *thing-thing* se concentra na verificação de problemas entre as “coisas” entre si e a Internet. Seus seis itens

estão organizados nas subcaracterísticas *de sensibilidade ao contexto e programabilidade*. Os itens descritos na categoria *human-thing* seguem uma organização similar, porém com a diferença de uma subcategoria a mais, *aspectos gerais da UX*. Os treze pontos de verificação apresentados na categoria *human-thing* são orientados em observar o nível de satisfação da experiência do usuário ao interagir no ambiente IoT. As categorias e subcategorias foram aspectos que ganharam destaque durante todo o processo de construção do *checklist* como descrito anteriormente na Seção 3.

Um fragmento do *checklist* é apresentado na Tabela 4 com categorias, subcategorias e quatro itens de verificação. O *checklist* completo pode ser acessado em: <https://goo.gl/xpq7w3>. Na Tabela 4 os itens quatro e cinco estão relacionados a identificação da utilidade, controle e compreensão do usuário sobre o ambiente IoT. Tais aspectos se demonstram primordiais para observação inicial, por conseguinte para a identificação da qualidade da UX em ambiente IoT. Os outros itens (14 e 19) se relacionam já ao atendimento das necessidades reais do usuário com o uso da aplicação. É importante ressaltar que o *checklist* está em processo de construção e o próximo passo no seu desenvolvimento é a aplicação em avaliações de UX no cenário IoT.

Tabela 4: Fragmento do Checklist de Observação da UX em cenário IoT

Categoria	Itens de Verificação	Essa situação ocorreu durante a avaliação do sistema IoT?			Tarefas em que ocorreu ?
		SIM	NÃO	N/A	
HUMAN-THING ASPECTOS GERAIS DA UX	4. O usuário consegue identificar corretamente todas as “coisas” que compõe o sistema IoT. (Ex.: Luzes, ares condicionados e travas de porta).”				
	5. O usuário consegue manipular corretamente as “coisas” para atingir seus objetivos na execução da tarefa.				
THING-THING SENS. AO CONTEXTO PROG.	14. O sistema se adequa corretamente às necessidades do usuário na execução da tarefa. (Ex.: O sistema oferece um serviço no momento que o usuário precisa.)				
	19. O sistema IoT não realizou as ações previamente programadas pelo usuário.				

4.1 Descrição de Uso da Ferramenta

O *checklist* precisa ser aplicado no mínimo por dois avaliadores, um responsável pela condução das tarefas a serem realizadas pelos usuários e outro avaliador dedicado a observação e preenchimento da ferramenta. Os avaliadores devem previamente verificar se a aplicação IoT apresenta as características de *sensibilidade ao contexto* e *programabilidade* e se todos os itens de verificação presentes do *checklist* serão observados na avaliação. Caso se identifique que algum item *não se aplica* ao escopo da avaliação, o observador deve marcar

previamente na ferramenta a opção “*não se aplica (N/A)*” para evitar distrações durante a observação. É recomendado que o observador estude a ferramenta e obtenha familiaridade com os itens de verificação. A realização da avaliação piloto pode auxiliar na familiaridade dos avaliadores com os pontos de verificação do *checklist*.

Recomenda-se o início do preenchimento da ferramenta pelos itens da categoria *human-thing* que estão relacionado ao contato direto do usuário com as “coisas” a fim de cumprir o conjunto de tarefas estabelecidas na avaliação. A subcategoria “aspectos gerais da UX” pode ser avaliada para todo o conjunto de tarefas do experimento. Já as subcategorias de “*sensibilidade ao contexto*” e “*programabilidade*” podem ser observadas em tarefas que propiciem sua observação em ambas as categorias. Dessa forma, recomenda-se que os avaliadores preparem o ambiente da avaliação para que consiga reproduzir situações contextuais chaves para o funcionamento da aplicação IoT, sejam elas realizadas em laboratório ou no contexto real de uso.

Após o preenchimento dos itens da primeira categoria, devem ser preenchidos os itens da *thing-thing*. Se sugere indicar no planejamento da avaliação a marcação das tarefas diretamente relacionadas a verificação desses itens. A reprodução das situações contextuais será primordial para a observação desses itens. É importante reforçar que os itens dessa categoria, se relacionam ao funcionamento do ambiente IoT e não são diretamente relacionadas as ações do usuário.

A interpretação do *checklist* pode ser realizada pela identificação dos itens mais recorrentes e críticos a UX. Após a delimitação da porção mais crítica, deve-se sugerir um conjunto de recomendações que auxiliem nas melhorias do sistema IoT, tendo em vista a melhoria da UX. Caso mais de um observador realize o preenchimento da ferramenta, os mesmos devem se reunir e consolidar os resultados da ferramenta.

Recomenda-se o uso da técnica *Thinking aloud* [7] para auxiliar na compreensão das atitudes e ações do usuário. A técnica precisa ser estimulada pelo condutor das tarefas do usuário e registrada pelo observador que deve realizar o preenchimento da ferramenta. Assim como algum sistema de *log* que auxilie na identificação dos problemas na *thing-thing*, visto que pode ser difícil ao avaliador perceber todos utilizando só observação.

5 CONCLUSÃO

Este trabalho apresentou a construção de um *checklist* para auxiliar avaliadores iniciantes a realizarem avaliação de UX no cenário de IoT. O *checklist* foi criado baseado na metodologia de Stufflebeam (2000) [13], onde foram criados os itens de verificação a partir dos resultados da revisão da literatura e de uma avaliação de UX no cenário de IoT. O artefato foi revisado e validado por 14 especialistas em IoT por meio de um questionário *online*. A versão atual da ferramenta apresenta um conjunto de 19 itens organizados nos tipos de interação presentes em um ambiente IoT. Como trabalho futuro pretende-se aplicar a ferramenta em avaliações de UX no cenário IoT. Com

isso, refinar o *checklist*, obtendo amadurecimento e evolução da ferramenta.

AGRADECIMENTOS

Rodrigo L. A. Almeida, Ticianne G. R. Darin, Italo L. de Araújo agradecem a FCPC pelas bolsas financiadas através de recursos da Lei de Informática e ao CNPq por patrocinar Rossana Maria de Castro Andrade com uma bolsa de produtividade DT-2, número de processo 314021/2009-4. Italo L. de Araújo agradece a CAPES pelo apoio recebido na sua pesquisa de doutorado.

REFERÊNCIAS

1. Rossana M.C. Andrade, Rainara M. Carvalho, Italo Linhares de Araújo, Káthia M. Oliveira, and Marcio E.F. Maia. 2017. What changes from ubiquitous computing to internet of things in interaction evaluation? In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 3–21. https://doi.org/10.1007/978-3-319-58697-7_1
2. Luigi Atzori, Antonio Iera, and Giacomo Morabito. 2010. The internet of things: A survey. *Computer networks* 54, 15: 2787–2805.
3. Johanna Bergman and Isabelle Johansson. 2017. The User Experience Perspective of Internet of Things Development.
4. Hao Cai, Egon Toft, Ole Hejlesen, John Hansen, Claus Oestergaard, and Birthe Dinesen. 2015. Health professionals' user experience of the intelligent bed in patients' homes. *International Journal of Technology Assessment in Health Care* 31, 4: 256–263. <https://doi.org/10.1017/S0266462315000380>
5. Elizabeth Sucupira Furtado, Gilberto Villa Nova, Isabela Gasparini, Leonardo Cunha de Miranda, Marcelo Barbosa, Tatiana Silva de Alencar, and Vânia Paula de Almeida Neris. 2017. Re-evaluating the First Five Years of the GrandIHC-Br-Challenge 3: Ubiquity, Multiple Devices and Tangibility. In *Proceedings of the XVI Brazilian Symposium on Human Factors in Computing Systems*, 69.
6. Jayavardhana Gubbi, Rajkumar Buyya, Slaven Marusic, and Marimuthu Palaniswami. 2013. Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems* 29, 7: 1645–1660. <https://doi.org/10.1016/j.future.2013.01.010>
7. D Hauser. 1989. Thinking aloud. *Heart Lung* 18, 3: 321–322. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=2722544
8. Lasse Kaila, Henrik Raula, Miika Valtonen, and Karri Palovuori. 2012. Living wood: a self-hiding calm user interface. In *Proceeding of the 16th International Academic MindTrek Conference*, 267–274.
9. Ibrahim Mashal, Osama Alsaryrah, Tein Yaw Chung, Cheng Zen Yang, Wen Hsing Kuo, and Dharma P. Agrawal. 2015. Choices for interaction with things on Internet and underlying issues. *Ad Hoc Networks* 28: 68–90. <https://doi.org/10.1016/j.adhoc.2014.12.006>
10. Alessandra Papetti, Andrea Capitanelli, Lorenzo Cavalieri, Silvia Ceccacci, Francesca Gullà, and Michele Germani. 2016. Consumers vs Internet of Things: A Systematic Evaluation Process to Drive Users in the Smart World. In *Procedia CIRP*, 541–546. <https://doi.org/10.1016/j.procir.2016.04.128>
11. Claire Rowland, Elizabeth Goodman, Martin Charlier, Alfred Lui, and Ann Light. 2015. *Designing Connected Products: UX for the Consumer Internet of Things*. O'Reilly Media, Inc.
12. Vagner F de Santana, Vânia Neris, Kamila R H Rodrigues, Renata Oliveira, and Newton Galindo Jr. 2017. Activity of Brazilian HCI Community from 2012 to 2017 in the Context of the Challenge'Future, Smart Cities, and Sustainability'. In *Proceedings of the XVI Brazilian Symposium on Human Factors in Computing Systems*, 67.
13. Daniel L Stufflebeam. 2000. Guidelines for developing evaluation checklists: the checklists development checklist (CDC). *Kalamazoo, MI: The Evaluation Center*. Retrieved on January 16, July: 2008. Retrieved from http://www.wmich.edu/evalctr/archive_checklists/guidelines_cdc.pdf%5Cnpapers3://publication/uuid/293B1EC6-6CA7-4CD1-9E80-885BC85B06A4
14. Jean-Philippe Vasseur and Adam Dunkels. 2010. *Interconnecting Smart Objects with IP: The Next Internet (Google eBook)*. Morgan Kaufmann. Retrieved from <http://books.google.com/books?hl=en&lr=&id=mVji-YA5kgEC&pgis=1>
15. Ovidiu Vermesan, Peter Friess, Patrick Guillemin, Sergio Gusmeroli, Sundmaeker Harald, Alessandro Bassi, Ignacio Soler Jubert, Margaretha Mazura, Mark Harrison, Markus Eisenhauer, and Pat Doody. 2011. Internet of Things Strategic Research Roadmap. *IoT European Research Cluster*, 9–52.

OpenData Processor: An Automation tool for the process of extracting and publishing open data to CKAN

Allyson Vilela and André Almeida
Federal Institute of Education, Science and Technology of
Rio Grande do Norte
Natal, Rio Grande do Norte, Brazil
{allyson.barros, andre.almeida}@ufrn.edu.br

Frederico Lopes
Metropole Digital Institute
Federal University of Rio Grande do Norte
Natal, Rio Grande do Norte, Brazil
fred@imd.ufrn.br

ABSTRACT

Public access to government information is an important aspect of modern society that allows an active participation of the population in monitoring government actions. Decree No. 8.777, signed on May 11, 2016, establishes the Open Data Policy of the Brazilian Federal Government. From this, the entities of the federal public administration, autarchic and foundational are obliged to make data available in open format. However, many of these institutions are failing to meet the commitments set out in the Decree. One possible explanation for this low number is the need for the technical team to have a good knowledge of their information systems and current legislation, allied to the difficulty of extracting the data, since in most institutions the whole process of data extraction, processing and publication of open data is done manually. In this sense, this work presents the OpenData Processor, an automation tool for the process of extracting, publishing and updating open data that brings agility in the publication and periodical updating, saving time and facilitating the management of open data portals.

KEYWORDS

open data. extraction tool. automated process. CKAN.

1 INTRODUÇÃO

O acesso à informação é um importante aspecto da sociedade moderna, uma vez que permite uma participação ativa da população na fiscalização das ações governamentais e contribui com a melhoria da gestão pública e no combate à corrupção [7]. Para o Tribunal de Contas da União, a abertura dos dados na Administração Pública contribui para uma melhor transparência na gestão pública a partir da contribuição da sociedade através da criação de soluções inovadoras para fiscalização desses dados [7]. Nesse sentido, o Decreto nº 8777 [1], assinado pela Presidência da República em 11 de Maio de 2016, institui a Política de Dados Abertos do Poder Executivo Federal.

De acordo com o *Open Knowledge International* [5], uma organização global sem fins lucrativos focada em difundir a importância do acesso e uso dos dados abertos como na sociedade civil, "dados são abertos quando qualquer pessoa pode livremente acessá-los, utilizá-los, modificá-los e compartilhá-los para qualquer finalidade, estando sujeito a, no máximo, a exigências que visem preservar sua

proveniência e sua abertura"[7]. Seguindo essa linha, dados abertos governamentais são aqueles gerados pelo governo e colocados à disposição da sociedade com o objetivo não disponibilizar apenas a sua leitura como também o seu acompanhamento, reutilização em sites e aplicações externas e o cruzamento com outros dados de órgãos/esferas diferentes [4].

Diante disso, os órgãos e entidades da administração pública federal direta, autárquica e fundacional são obrigados a publicarem os seus dados de forma aberta, segundo as diretrizes definidas na Infraestrutura Nacional de Dados Abertos (INDA) do Ministério do Planejamento, Orçamento e Gestão (MPOG) [3].

Este artigo tem por objetivo apresentar as principais características do OpenData Processor, uma aplicação web que automatiza todo o processo de extração, tratamento e catalogação de dados abertos seguindo os padrões definidos pela Infraestrutura Nacional de Dados Abertos (INDA) do Ministério do Planejamento, Desenvolvimento e Gestão (MPOG) e que pode ser utilizada por instituições públicas dos governos federal, estaduais e municipais. Essa versão do OpenData Processor utiliza o CKAN como plataforma de dados abertos. Tal plataforma é mais utilizada no mundo para esse fim. De código livre e desenvolvida em Python, a CKAN é utilizada no Portal Brasileiro de Dados Abertos e por governos e organizações públicas de diversos países [6].

2 PROBLEMÁTICA E MOTIVAÇÃO

O Decreto nº 8777, assinado pela Presidência da República em 11 de Maio de 2016, institui a Política de Dados Abertos do Poder Executivo federal. A partir dele, os órgãos e entidades da administração pública federal direta, autárquica e fundacional são obrigados a disponibilizarem os dados, que não estejam sob sigilo ou sob restrição de acesso, contidos em seus banco de dados sob a forma de dados abertos possibilitando assim um aprimoramento da cultura de transparência pública para a sociedade e facilitar a troca de informações entre órgãos e entidades da administração pública federal e as diferentes esferas da federação [1].

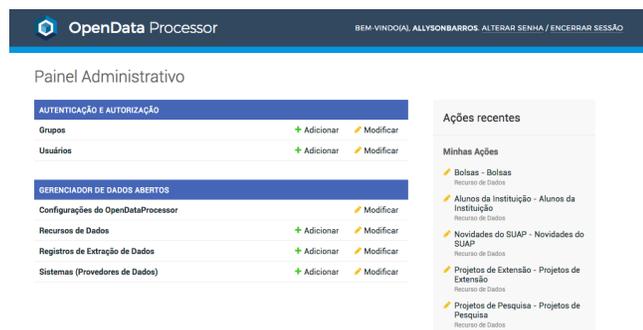
No entanto, apesar dessa obrigação, grande parte das instituições do poder executivo federal estão descumprindo os compromissos definidos no Decreto [2]. Uma possível explicação para esse número é a dificuldade enfrentada durante o processo de abertura dos dados devido a falta de integração entre os sistemas de informação utilizados nas instituições e as plataformas de dados abertos.

Além disso, em muitas instituições esse processo é feito de forma manual e exige que a equipe responsável pela abertura dos dados possuam bons conhecimentos de seus sistemas de informação e das legislações e recomendações vigentes.

In: XVII Workshop de Ferramentas e Aplicações (WFA 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pôsteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.

© 2018 SBC – Sociedade Brasileira de Computação.
ISBN 978-85-7669-435-9.

Figura 1: Captura de tela do painel administrativo do OpenData Processor.



Nesse sentido, o desenvolvimento de uma ferramenta de automação do processo de extração, publicação e atualização dos dados traz uma maior agilidade na publicação e atualização periódica dos dados abertos, economia de tempo e facilidade na gestão dos portais de dados abertos das instituições públicas.

3 FUNCIONALIDADES

O objetivo do OpenData Processor é automatizar todo o processo de extração, publicação e atualização de dados abertos das instituições públicas seguindo os padrões definidos pela Infraestrutura Nacional de Dados Abertos (INDA) do Ministério do Planejamento, Desenvolvimento e Gestão (MPDG).

A ferramenta possui diversas funcionalidades que visam facilitar o gerenciamento dos dados abertos das instituições e tornar a publicação dos dados abertos mais fácil por parte da equipe de TI. A seguir serão detalhadas algumas dessas funcionalidades:

- **Painel Administrativo:** Através da interface administrativa é possível realizar o cadastro, edição ou exclusão de organizações, grupos, conjuntos de dados e executar as tarefas de extração e sincronização de dados. A Figura 1 apresenta a interface administrativa do OpenData Processor.
- **Sincronização de dados com o CKAN:** O OpenData Processor permite a integração com instâncias já existentes do CKAN. Com isso, após cadastrar ou atualizar as configurações relativas à integração com o CKAN, é realizada de forma automática o cadastro/atualização das Organizações, Grupos e Conjuntos de Dados existentes na instância CKAN da Instituição.
- **Suporte à múltiplos Provedores de Dados:** O OpenData Processor permite a extração de dados a partir de múltiplos provedores de dados diferentes: sistemas de informação, banco de dados e planilhas eletrônicas.
- **Extração de Dados via API REST:** Permite a extração de dados a partir de uma API REST já existente que utilize um dos seguintes mecanismos de autenticação: Nenhuma, HTTP Basic, Token Simples, OAUTH ou JSON Web Token (JWT).
- **Extração de Dados via Banco de Dados:** Permite a extração de dados a partir dos SGDBs mais utilizados no

Figura 2: Captura de tela da pré-visualização dos dados à serem extraídos.

Amostra dos Dados Extraídos							
status	valor_liquido_contabil	campus	valor_inicial	estado_conservacao	codigo	id	descricao
ativo	0,01	CNAT	0,01	Não aplicar	2	18070	POLTRONA GIRATORIA MOD. C-1, C/ ASSENTO E ENCOSTODE MADEIRA.
baixado	0,01	CNAT	0,01	Não aplicar	8	23827	FICHARIO DE MADEIRA, C/11 GAVETAS, MED.0,84X0,53X0,28M.
baixado	0,01	CNAT	0,01	Não aplicar	21	18071	MESA, MOD.MM-2, C/01 GAVETA, MED.0,55X0,45X0,70M,C/RODIZIOS.
baixado	0,01	CNAT	0,01	Não aplicar	22	16494	FICHARIO DE ACO, MARCA "CONFIANCA", C/02 GAVETAS, P/FICHAS DE 5X8, MOD. 0258, DE COR CINZA.
baixado	0,01	CNAT	0,09	Não aplicar	40	121	ESTANTE DE ACO, C/06 PRATELEIRAS, MED. 0,92X0,30X1,98M, DE COR CINZA.
baixado	0,01	CNAT	0,09	Não aplicar	45	122	ESTANTE DE ACO, C/06 PRATELEIRAS, MED. 0,92X0,30X1,98M, DE COR CINZA.
baixado	0,01	CNAT	0,09	Não aplicar	46	123	ESTANTE DE ACO, C/06 PRATELEIRAS, MED. 0,92X0,30X1,98M, DE COR CINZA.

Figura 3: Captura de tela da configuração do agendamento de extração automática no OpenData Processor.

mercado como SQLite3, MySQL / MariaDB, PostgreSQL e SQL Server e também a partir de bancos de dados NoSQL como Redis, Memcached, MongoDB, Riak e CouchDB.

- **Pré-visualização dos dados:** Permite a pré-visualização dos dados à serem extraídos via API ou Banco de Dados. Dessa forma, é possível validar se as configurações relativas à extração foram realizadas corretamente. A Figura 2 ilustra a configuração do agendamento de extração automática.
- **Dicionário de Dados:** Permite a criação de uma página de especificação, denominada dicionário de dados, que contém o nome, tipo, formato e descrição de cada campo do conjunto de dados.
- **Agendamento de Extração Automática de Dados:** Ao cadastrar ou editar um Conjunto de dados é possível indicar se a extração automática será executada diariamente, semanalmente, mensalmente, semestralmente ou anualmente e em qual horário ela será executada. Ao salvar as alterações, o agendamento da extração é adicionado à fila de execução em segundo-plano. A Figura 3 ilustra a configuração do agendamento de extração automática.
- **Auditoria:** A ferramenta permite que os administradores possam visualizar as ações de cadastro, edição e exclusão realizadas pelos usuários do sistema a partir da interface administrativa.

Figura 4: Captura de tela de acompanhamento das extrações realizadas pelo OpenData Processor.

Todas as datas: Julho 2018

Ação: 0 de 35 selecionados

RECURSO - CONJUNTO DE DADOS	DATA E HORÁRIO DE INÍCIO DA EXTRAÇÃO	DATA E HORÁRIO DE TÉRMINO DA EXTRAÇÃO	QUANTIDADE DE REGISTROS
Planos Individuais de Trabalho - Planos Individuais de Trabalho	12/07/2018 00:30:51	12/07/2018 00:31:13	1212
Contratos de 2015 - Contratos	11/07/2018 10:07:45	11/07/2018 10:10:42	206
Contratos de 2016 - Contratos	11/07/2018 09:58:02	11/07/2018 09:58:15	148
Contratos de 2017 - Contratos	11/07/2018 09:57:47	11/07/2018 09:58:02	212
Contratos de 2018 - Contratos	11/07/2018 09:57:39	11/07/2018 09:57:47	163
Processos de 2015 - Protocolo	11/07/2018 08:40:13	11/07/2018 09:01:13	50604
Processos de 2016 - Protocolo	11/07/2018 08:31:39	11/07/2018 08:40:12	33552
Processos de 2017 - Protocolo	11/07/2018 08:07:06	11/07/2018 08:31:39	62207
Processos de 2018 - Protocolo	11/07/2018 07:58:38	11/07/2018 08:07:06	33346
Projetos de Pesquisa - Projetos de Pesquisa	08/07/2018 03:30:10	08/07/2018 03:34:36	2711
Setores - Setores	08/07/2018 02:30:07	08/07/2018 02:30:40	597
Projetos de Extensão - Projetos de Extensão	07/07/2018 03:00:59	07/07/2018 03:05:51	2626

- **Acompanhamento de Extrações de Dados Realizadas:** Ao executar uma extração de dados, independentemente do resultado, o OpenData Processor registra as informações relativas à execução tais como se a execução foi bem sucedida ou não, a quantidade de registros que foram extraídos e o horário de início e término da extração. A Figura 4 ilustra a funcionalidade de acompanhamento das extrações realizadas.

4 ARQUITETURA E DEPLOY

O OpenData Processor é uma solução web escalável que foi desenvolvida utilizando as seguintes tecnologias: Python, Django, PostgreSQL, Redis, Python-RQ, Pandas, Docker e Docker Compose.

A ferramenta se adapta aos diferentes tipos de infraestrutura de TI disponíveis nas instituições públicas, ou seja, a partir de três opções arquiteturais disponíveis é possível implantá-la na infraestrutura de TI da instituição, em um provedor de nuvem como AWS e Azure, ou utilizá-la como um serviço (SaaS).

Nas Seções 4.1, 4.2 e 4.3 serão apresentadas cada uma dessas opções arquiteturais e na Seção 4.4 serão apresentadas informações relativas ao *deploy* do OpenData Processor.

4.1 Standalone

Esta opção arquitetural permite que o OpenData Processor seja executado em um ambiente completamente isolado do ambiente de produção dos sistemas corporativos utilizados na instituição. Dessa forma, todo o processo de extração dos dados é realizado em um único servidor virtual, evitando que esse processo interfira no desempenho dos sistemas da instituição. A figura 5 ilustra a separação dos componentes do OpenData Processor na versão "Standalone".

Os componentes dessa arquitetura são: **Queue**, responsável pelo gerenciamento da fila de execução de tarefas em *background*; **Scheduler**, responsável pelo agendamento de tarefas na fila de execução em *background* e **Worker**, responsável pela execução das tarefas em *background*.

A comunicação entre o OpenData Processor, o CKAN e os sistemas de informação utilizados nas instituições é feita através de

Figura 5: Visão geral dos componentes da arquitetura "Standalone".

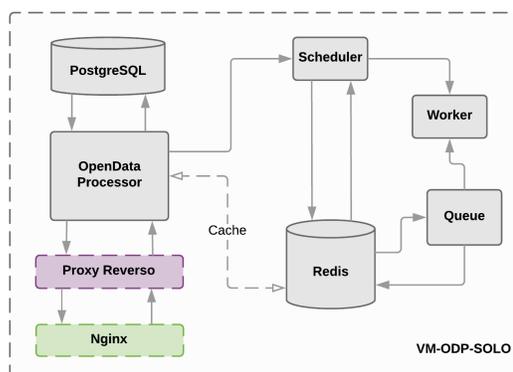
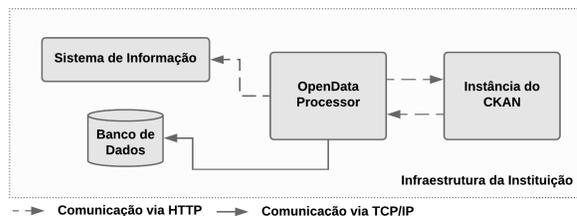


Figura 6: Ilustração da forma de comunicação entre o OpenData Processor, os provedores de dados e o CKAN na versão Standalone.



serviços web REST, já a comunicação com os bancos de dados são feita através de conexão TCP/IP. A Figura 6 ilustra a forma de comunicação entre o OpenData Processor, os provedores de dados (sistemas de informação ou banco de dados) e o CKAN na versão da arquitetura Standalone.

4.2 Distribuída

Esta opção de arquitetura é ideal para as instituições que possuem uma infraestrutura de TI mais robusta ou que utilizam provedores de serviços de hospedagem na nuvem como AWS, Azure, DigitalOcean e Scaleway uma vez que a arquitetura permite a separação de cada um de seus componentes em servidores virtuais separados ou através de tecnologias baseadas em contêineres Linux (*Linux Containers - LXC*) como Docker, por exemplo.

Ao utilizar essa opção arquitetural é possível distribuir o consumo de recursos computacionais, uma vez que ela permite a paralelização das rotinas de extração de dados através da criação de novas instâncias do componente Worker. A figura 7 ilustra a separação dos componentes da arquitetura em uma infraestrutura baseada em LXC.

A comunicação entre o OpenData Processor, o CKAN e os provedores de dados se dá da mesma forma que na versão "Standalone", no entanto, caso o OpenData Processor seja implantado em um provedor de Nuvem, é necessário a liberação de conexões externas aos servidores de banco de dados utilizados. A Figura 8 ilustra a

Figura 7: Visão geral dos componentes da arquitetura "Distribuída" enfatizando a separação dos componentes em uma Infraestrutura Baseada em LXC.

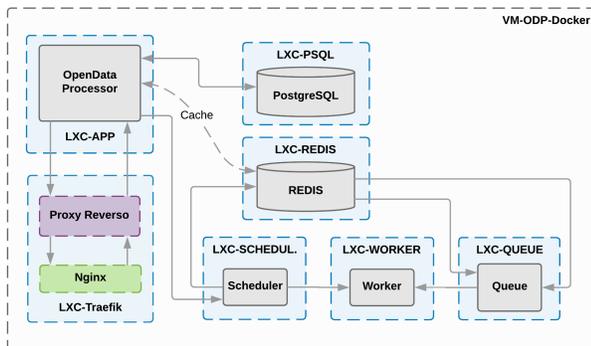
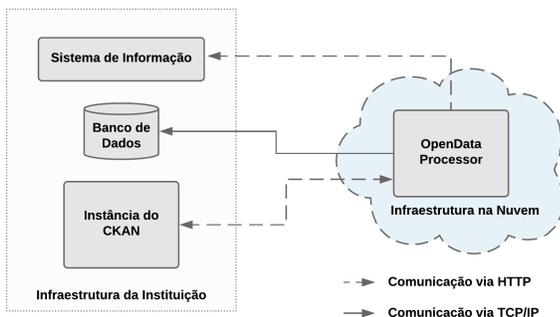


Figura 8: Ilustração da forma de comunicação entre o OpenData Processor, os provedores de dados e o CKAN na versão "Distribuída" implantada em um provedor de Nuvem.



forma de comunicação entre o OpenData Processor, os provedores de dados (sistemas de informação ou banco de dados) e o CKAN na versão da arquitetura Distribuída.

4.3 Software como Serviço (SaaS)

Esta opção arquitetural é ideal para as instituições que não possuem infraestrutura de TI e/ou equipe técnica para manter o OpenData Processor funcionando. Ou seja, por se tratar de um software como serviço, o fornecedor do software se responsabiliza por toda a estrutura necessária à disponibilização do sistema.

Nesta versão da arquitetura foi adicionado um novo componente, denominado "gerenciador de instâncias", o qual é responsável pela criação, exclusão, desativação e ativação das instâncias do OpenData Processor. Além disso, ele permite que as instituições solicitem a criação de uma instância de avaliação por um período de até 60 dias a partir do website¹ de divulgação da ferramenta.

Vale salientar que, as instâncias compartilham tanto a infraestrutura física de servidores quanto o banco de dados uma vez que

¹<http://www.opendataprocessor.com>

Figura 9: Visão geral dos componentes da arquitetura "Software como Serviço".

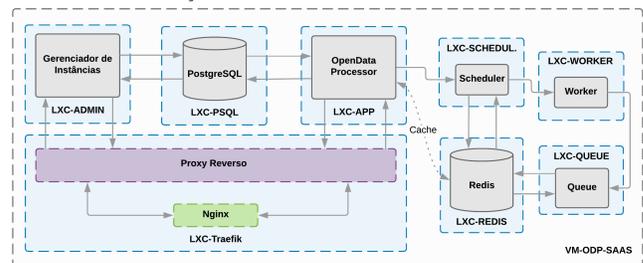
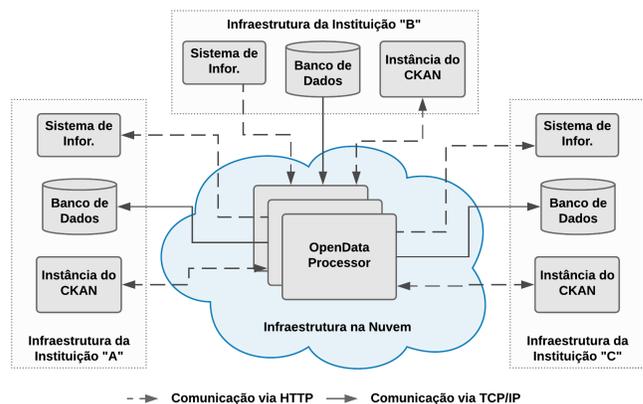


Figura 10: Ilustração da forma de comunicação entre o OpenData Processor, os provedores de dados e o CKAN na versão SaaS.



a separação dos dados é feita de forma lógica através de *schemas* do SGDB. Dessa forma, ao criarmos quatro novas instâncias serão criados quatro *schemas* no banco de dados, cada um referente a uma instância.

Da mesma forma que na versão "Distribuída", como a instância é executada fora da rede da instituição é necessário a liberação de conexões externas aos servidores de banco de dados utilizados como provedores de dados. A Figura 8 ilustra a forma de comunicação entre o OpenData Processor, os provedores de dados (sistemas de informação ou banco de dados) e o CKAN na versão da arquitetura Software como Serviço.

4.4 Deploy

Como visto anteriormente, as instituições podem optar por realizar a implantação do OpenData Processor em sua própria infraestrutura de TI, em um provedor de nuvem ou utilizá-lo como um serviço. Dessa forma, para implantar o OpenData Processor são necessárias as seguintes configurações mínimas, dependendo da opção arquitetural escolhida:

- **Standalone:** Um servidor virtual com 2 processadores, 2GB de Memória de RAM e 20GB de espaço em disco utilizando a distribuição Linux Debian ou Ubuntu Server. No

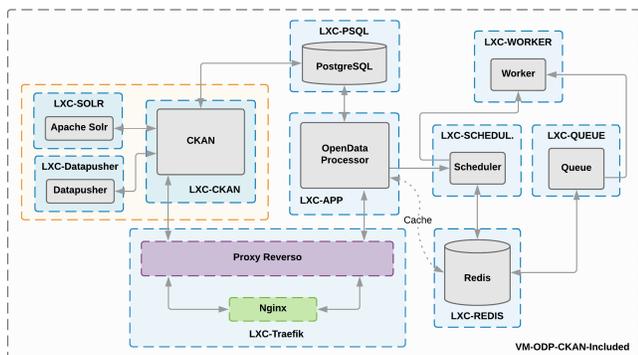


Figura 11: Visão geral da arquitetura do OpenData Processor enfatizando a inclusão dos componentes da versão customizada do CKAN.

entanto, caso a instituição deseje publicar conjuntos de dados contendo milhões de registros, será necessário a utilização de, no mínimo, 4GB de memória RAM.

- **Distribuída:** Pelo menos três servidores virtuais com 2GB de Memória de RAM e 20GB de espaço em disco utilizando a distribuição Linux Debian ou Ubuntu Server, sendo um para o banco de dados, outro para a execução das tarefas em *background* e outro para o servidor de aplicação. Ou, caso deseje executar os componentes em contêineres Linux, um servidor virtual com 4 processadores, 8GB de Memória de RAM e 40GB de espaço em disco utilizando a distribuição Linux Debian ou Ubuntu Server.
- **Software como Serviço:** Como a infraestrutura necessária para a implantação da instância é de responsabilidade do fornecedor do Software, a instituição não precisa se preocupar com as configurações mínimas.

Para facilitar a implantação do OpenData Processor nas instituições, para cada uma das versões arquiteturais são disponibilizados scripts que automatizam o processo de *deploy* através da utilização do Docker e Docker Compose. Dessa forma, todo o processo de instalação e configuração das dependências necessárias para o funcionamento do OpenData Processor é feito em poucos minutos sem a necessidade da intervenção humana.

Além disso, visando facilitar o processo de abertura dos dados em instituições públicas que ainda estão em fase de planejamento da abertura de seus dados, o OpenData Processor possibilita o *deploy* automático de uma instância do CKAN. Essa funcionalidade evita que a equipe de TI perca tempo com a instalação e configuração do CKAN. A Figura 11 ilustra a adição dos componentes do CKAN, destacados em amarelo, aos componentes da versão distribuída do OpenData Processor.

Os scripts de automatização do *deploy* do CKAN foram baseados em um projeto² de código aberto oferecido pela Plataforma FIWARE³ e está disponível para utilização no endereço <https://github.com/allysonbarros/docker-opendaprocessor-ckan>. A Figura 12 ilustra o visual da instância customizada do CKAN.

²<https://github.com/okfn/docker-fiware-ckan>

³<https://www.fiware.org/>

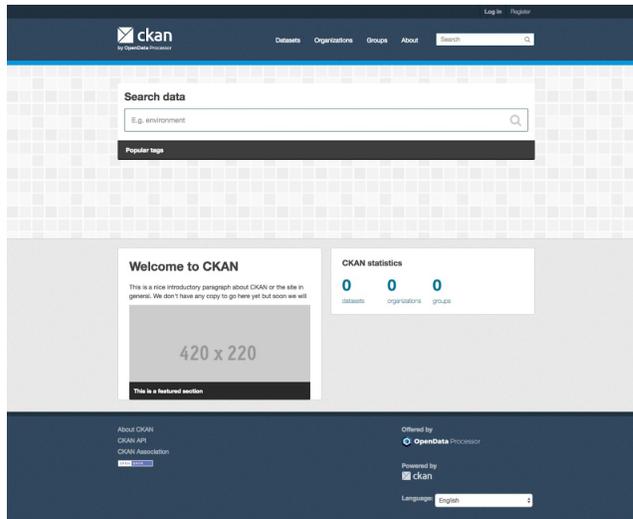


Figura 12: Captura de tela da página inicial da instância customizada do CKAN oferecida pelo OpenData Processor.

5 LICENÇA

O OpenData Processor é um software proprietário que foi registrado junto ao Instituto Nacional da Propriedade Industrial (INPI) e sua titularidade é compartilhada entre o IFRN e a UFRN.

Dessa forma, o processo de licenciamento para a utilização do OpenData Processor pelas instituições públicas se dá de duas formas:

- Através da assinatura de convênio de cooperação técnica para a realização da transferência de tecnologia entre a instituição e os titulares do software;
- Através da contratação de uma das empresas licenciadas pelos titulares para a realização da implantação, customização e suporte técnico do software.

REFERÊNCIAS

- [1] BRASIL. 2016. Decreto nº8777. (2016). http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/D8777.htm
- [2] CONTROLADORIA-GERAL DA UNIÃO. 2018. Painel Monitoramento de Dados Abertos. (2018). <http://paineis.cgu.gov.br/dadosabertos/index.htm>
- [3] MINISTÉRIO DO PLANEJAMENTO, DESENVOLVIMENTO E GESTÃO. 2017. Plano de Dados Abertos. (2017). <http://wiki.dados.gov.br/Plano-de-Dados-Abertos.ashx>
- [4] NÚCLEO DE INFORMAÇÃO E COORDENAÇÃO DO PONTO BR. 2011. Manual dos dados abertos: Governos. (2011). http://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf
- [5] OPEN KNOWLEDGE INTERNATIONAL. 2017. The Open Definition. (2017). <http://opendefinition.org/>
- [6] OPEN KNOWLEDGE INTERNATIONAL. 2018. CKAN. (2018). <http://ckan.org/about/>
- [7] TRIBUNAL DE CONTAS DA UNIÃO. 2016. Cinco Motivos para a Abertura de Dados na Administração Pública. (2016). <http://portal.tcu.gov.br/lumis/portal/file/fileDownload.jsp?fileId=8A8182A24F0A728E014F0B366F2E2A40>

MobSink: a Visual Mobile Wireless Sensor Networks Positioning Simulator

João Paulo Just Peixoto
Federal Institute of Bahia
Vereador Romeu Agrário Martins St., Tendo
Valença, BA 45400-000
joao.just@ifba.edu.br

Daniel G. Costa
State University of Feira de Santana
Transnordestina Av., Novo Horizonte
Feira de Santana, BA 44036-900
danielgcosta@uefs.br

ABSTRACT

The planning and deployment of a WSN (Wireless Sensor Network) in a Smart City can be a very challenging work. People involved in such implementation must be aware of how sensors and sinks may behave in the target environment. To aid the development of new WSNs, simulation tools are often used to predict how nodes will interact before spending time and money in a real deployment. Also, simulator can help students better understand WSNs. In this paper, we present MobSink, a simulator for WSNs with multiple mobile sinks. MobSink also allows to perform simulations in a Smart City scenario, with streets and movements constraints. It also shows how to configure MobSink for a generic scenario and describes how it works internally.

KEYWORDS

simulation, WSN, sinks mobility, smart city

1 INTRODUCTION

A Wireless Sensor Network (WSN) is an ad-hoc network composed of several sensor nodes that gather environment data as temperature, humidity, luminosity and also multimedia data as audio and video. The sensed data is then transmitted in a hop-by-hop manner to centralized nodes called *sinks*, which can be connected to the Internet or to a local server to deliver all the collected data to an application [9].

The main purpose of WSNs is to monitor environments, applying the data obtained from the network in specific applications [1]. Sensor nodes can also be deployed in a Smart City to collect many useful information as pollution status, urban noise, local temperature, traffic density, etc. There are several possibilities of applications of WSNs in Smart Cities, like smart surveillance, smart transportation, smart services, among others [3]. In such scenarios, sinks can be installed in vehicles to run across the roads, collecting data from sensor nodes as they pass by them [6].

In an ordinary WSN, the sensor nodes have some hardware constraints: restricted energy supply, reduced processing power, low memory, low communication range, etc. Most of the constraints of a sensor node are related to the low power available (most sensor nodes run with non-rechargeable batteries [2]).

In order to create a sensing environment, several sensor nodes must be deployed into the area of interest, which can be time consuming and costly if the network is not well designed [2]. When trying to predict the performance of a WSN before its deployment, simulators can be used to support better designing of the network [4, 7, 8]. Also, they can help students to learn more about WSN configurations without the need to buy and deploy real hardware.

In this paper we describe the development of the new tool MobSink, a software specially designed to simulate WSNs in a Smart City environment. The main purpose of MobSink is to test energy consumption and data delivery in a WSN with mobile sinks in a Smart City. A traffic model is used to calculate the fastest route of a mobile sink in such a way it can avoid traffic jams and reach its destination as fast as possible. The use of this type of simulator can aid the designing and deployment of WSNs in a Smart City scenario, as well to help students to better understand how a WSN works with multiple mobile sinks. MobSink is focused in multiple sinks positioning and do not consider routing algorithms, transport protocols, etc.

MobSink is licensed under GNU GPLv3 (GNU General Public License version 3) and its source code is available on GitHub. This makes possible to students to learn how to code tools for WSN simulation and also contribute with this project.

The remainder of this paper is organized as follows. Section 2 presents the energy model and positioning scheme used in MobSink, while section 3 explains how these models were implemented; section 4 describes how to simulate a Smart City scenario and 5 concludes this paper and presents future works.

2 MOBSINK INTERNALS

The MobSink tool works by simulating sensor nodes that generate data and that forward the buffered data to the next hop, which is another sensor node, until the data reaches the nearest sink. Actually, there are different possible behaviors for the sensor nodes, but the MobSink tool was designed to incorporate the relevance-based scheme proposed in [5]. In that proposal, each sensor node has a *RL* (Relevance Level) that varies from 0 (no relevance) to 15 (highest relevance). Generally speaking, the relevance level of each sensor node can be defined in different ways, but the values of *RL* will usually be a reflection of the applications monitoring requirements (e.g. a temperature sensor near a nuclear reactor may have higher relevance than a sensor node monitoring the ambient temperature in the same WSN). It is reasonable to expect that the amount of data generated by a sensor node depends on its *RL* (*RL* = 0 means no data is generated by the sensor and it works only as a relay node).

In: XVII Workshop de Ferramentas e Aplicações (WEA 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pôsteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.
© 2018 SBC – Sociedade Brasileira de Computação.
ISBN 978-85-7669-435-9.

In wireless sensor networks, data transmission and relaying tasks are expected to consume most energies of the nodes. Therefore, the position of the sinks influence the path data packets must follow to reach their destinations, also impacting in the amount of consumed energy. This particularity of WSNs was specially considered when designing the MobSink tool.

The following subsections describe how MobSink computes the energy consumed by each sensor node and how this tool moves the sink nodes inside the modelled WSN.

2.1 Energy consumption

The MobSink tool operates generating data packets at source nodes and transmitting them to the nearest sink, which is performed in the scope of a “simulation”. At each iteration of a simulation, every sensor node generates PDUs (Packet Data Units) according to its RL (sensor nodes with $RL = 0$ generate no data). Each generated packet is then inserted into a buffer for further transmission. If the buffer is full, the exceeding packets are discarded and this information is accounted for a final report. After inserting the generated PDUs into the local buffer, the sensor node transmits to the next hop, which receives the packets and stores them into its buffer (again, if the receiver node’s buffer is full, the packets are dropped).

This whole process is performed once at each iteration for every sensor node in the WSN simulation. Data transmission and reception procedures result in energy consumption in each node that participated in the considered communication. Currently, MobSink uses the power scheme described in [5]: each PDU is 127 bytes sized, the power used in transmission and reception of packets is 60 mW and 4 μ s are needed to transmit each bit of data (computed according to the maximum transmission rate, which is based on the IEEE 802.15.4 standard). Also, each sensor node is supplied by a 3 V battery with 1000 mAh capacity and has a 128 KB buffer. Although the current version of MobSink does not allow for configurable power source and buffer, future releases will feature a configuration file.

When a packet leaves the source node, it is relayed by some intermediate nodes until it reaches the sink. The energy consumed in each intermediate node is computed to account the total energy spent in the simulation. Let $D_{tx}(i)$ be the size of the packet transmitted from node i in bits, $P_{tx}(i)$ the power needed by node i to transmit a bit and $t(i)$ the time needed by node i to transmit a bit, $E(i)$ is then the total power needed by node i to transmit a packet, as shown in Equation 1.

$$E_{tx}(i) = D_{tx}(i) \cdot P_{tx}(i) \cdot t(i) \quad (1)$$

In the same way, MobSink computes the energy needed to receive a packet in each sensor node. The equation is almost the same, but MobSink considers the size of the received packet in node i as $D_{rx}(i)$ and the reception power as $P_{rx}(i)$ in Equation 2.

$$E_{rx}(i) = D_{rx}(i) \cdot P_{rx}(i) \cdot t(i) \quad (2)$$

In a sequence $S = (s_0, s_1, \dots, s_n)$ of sensor nodes, the total energy spent to send a packet from its source to the sink is obtained by summing the energy spent by each node to transmit the packet to the next hop and the energy spent by each node to receive the packet from its origin, as described in Equation 3.

$$\sum_{i=0}^{n-1} E_{tx}(i) + E_{rx}(i + 1) \quad (3)$$

2.2 Sinks positioning

MobSink provides three main positioning schemes for sinks: static, fixed movement and relevance-based.

In the static scheme, the sinks are positioned in a grid centralized manner and they keep their initial positions during the whole simulation time. No movements are performed in this scheme. Distant sensor nodes must relay its data through intermediate nodes to reach the sink. Although currently MobSink does not allow for a custom fixed position, this is a feature to come in a future release.

In the fixed movement scheme, the sinks move on straight lines which can be a horizontal or a vertical path through the WSN. The sink moves at a constant speed and just like the static scheme, its initial position is defined by the simulator.

At last, in the relevance-based scheme, MobSink uses the approach defined in [6]. In such way, the sinks positions are defined by the relevance levels of sensor nodes. The goal of this approach is to make sinks closer to more relevant nodes, potentially improving the global performance of the network, since sinks will be closer to the source nodes that transmit more data (higher values of RL). If there are more than one sink in the WSN, this algorithm creates clusters of sensors according to their relevances. The total RL of the clusters, defined by the sum of the sensors relevances in a cluster, are balanced in a way that each cluster total RL do not differ too much from the others.

When using the relevance-based scheme, a Smart City scenario can be easily simulated. The user can model streets of a city and the sinks will move through these streets to reach its final positions. Every time the topology changes (e.g.: a sensor node RL changes or runs out of energy), MobSink recalculates new positions for the sinks and make them move. The sinks paths always obey roads restrictions, if defined by the user.

3 SOFTWARE ARCHITECTURE

The MobSink was designed to model the operation of any WSN that has one or more static or moving sinks, which complies with a lot of different real sensor networks. For that, the operation details of wireless sensor networks were carefully considered.

This section explains how a WSN is simulated and how sinks positions are calculated in a Smart City scenario. Subsection 3.1 presents the mechanism used to calculate energy consumption and data traffic in a simulation, while subsection 3.2 presents the developed mechanisms to model a city and its associated traffic behavior.

3.1 WSN simulation

To simulate a WSN, some classes were implemented to represent the actors in a WSN. Figure 1 presents a class diagram showing the relations of nodes and sinks in MobSink. When the user inserts any new sensor node in the WSN, an object of “Node” class is instantiated. After specifying how many sinks will be used in the simulation, MobSink instantiates a “Sink” object for each sink in simulation and a “Cluster” object for each sink. Each “Cluster” object has a

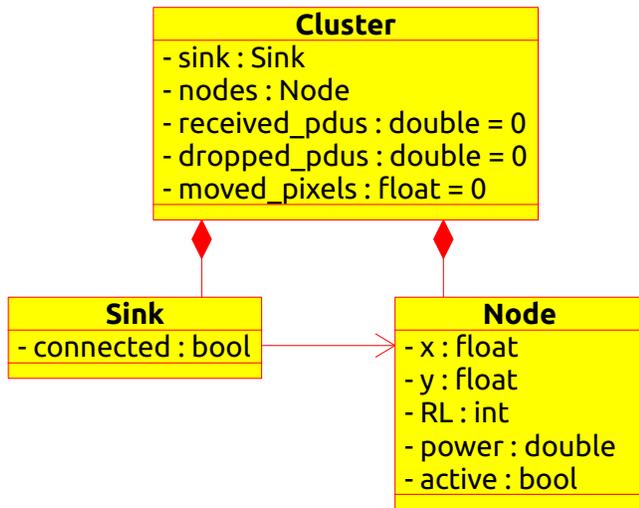


Figure 1: Class diagram for MobSink.

sink associated to it, besides several nodes. The nodes associations are made depending on the chosen positioning algorithm.

When simulation starts, MobSink positions every sink according to the employed positioning algorithm, associating each node to its nearest sink. This is done by adding a reference of the node object to the correspondent Cluster object. After this, MobSink has some clusters (one for each sink), containing its nodes. The sensor nodes can only communicate with other nodes in the same cluster, including the sink.

MobSink simulates “one second” intervals. At each iteration, it checks every sensor node and sends it a message to “work” (generate data and update its status according to user’s programmed schedule). The sensor node will generate packets according to its relevance level and sinks may be repositioned if there was any topology change in the network (e.g: a sensor toggles active/inactive or changes its relevance level). At last, active sensors with buffered data transmit to the next hop. At the end of the iteration, one second has passed in the simulated scenario and every active sensor has generated and/or relayed data. Figure 2 presents the flowchart of MobSink simulation.

After simulation finishes, the cluster object has accounted the total of received and dropped packets, distance traveled by its sink and power consumption of all sensor nodes.

3.2 Smart City modeling

In order to model a Smart City scenario, MobSink simulates roads and their traffic behavior. Before simulation begins, MobSink checks every road added by the user and finds all intersections between them. Each road is modeled as an edge in a directed weighted graph. Roads’ source and destination points are modeled as vertices and so are their intersections.

To set a weight in each edge, MobSink uses the total time to traverse the road, calculated using the road’s length and average speed as parameters. Our approach to simulate traffic jams was to use the roads legal speed limit defined as $V(r)$ as a maximum

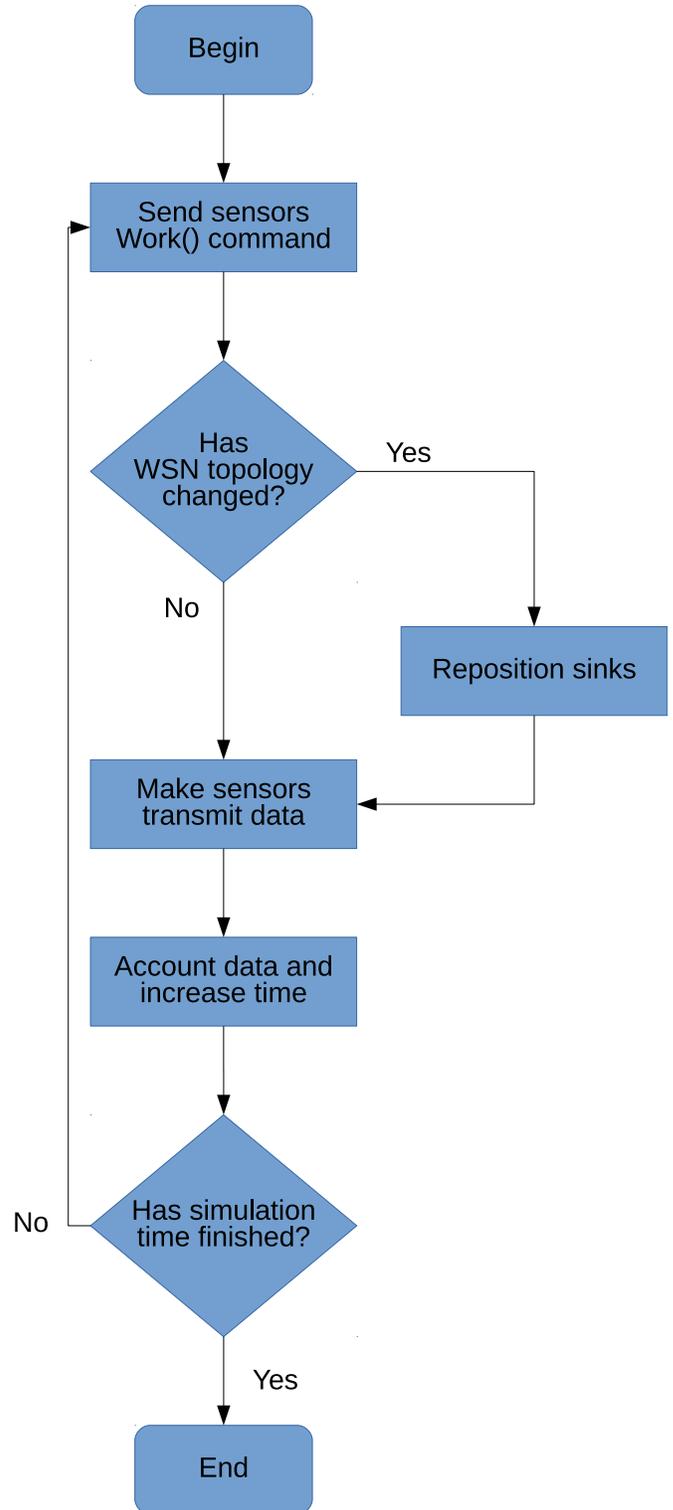


Figure 2: Flowchart of MobSink simulation.

vehicle speed, applying then a traffic coefficient $m(r)$ (as shown in Equation 4) to set the approximated time to traverse each road. By using these parameters with the road's length, defined as $S(r)$ in that Equation, it is possible to calculate the time spent to traverse a road r . The user can then program a traffic schedule and simulate traffic jams along the simulation time.

$$t(r) = \lim_{x \rightarrow m(r)} \frac{S(r)}{V(r) \cdot x} \quad (4)$$

Every time a sink has to be positioned in the map, Dijkstra's algorithm is run, using the specified weight ($t(r)$) to find the fastest route, avoiding traffic jams. The simulator also offers an option to use the shortest routes, ignoring traffic information. A control schedule is implemented in a WSN XML file, setting traffic data for each road.

3.3 WSN XML programming

Sensors behavior and traffic can be dynamically changed during the simulation. The way MobSink allows it to be done is by specifying the changes and the times they occur in a well-formatted WSN XML file. Each sensor node is defined by a *sensor* tag in the XML file, containing its coordinates in the map and the initial relevance level. Inside the *sensor* tag, a *control* tag is used to define a change in the sensor behavior. The *control* tag has three attributes: *time*, to specify at which time of simulation the change will be applied; *rl*, to specify the new relevance level of the sensor; and *enabled*, to enable or disable the sensor node at that time of simulation.

The roads, if any, are defined by a *path* tag, containing its two vertices coordinates and flow (it may be an one-way road). Like the sensor nodes, the roads may have *control* tags with three attributes: *time*, like sensor nodes, to specify the time of change; *speedlimit*, to specify the new speed limit from that time on; and *traffic*, varying from 0 (stopped traffic) to 1 (free road), to specify the traffic behavior at that moment. The *traffic* attribute is used as $m(r)$ in Equation 4. Listing 1 shows an example of a WSN XML file for MobSink.

Listing 1: MobSink XML sample

```
<?xml version="1.0" encoding="UTF-8"?>
<network width="200" height="200">
  <path xa="0" ya="50" xb="100" yb="50">
    <traffic time="1" speedlimit="60"
      traffic="1"/>
    <traffic time="3600" speedlimit="60"
      traffic="0.3"/>
  </path>
  <sensor x="10" y="40" rl="1">
    <control time="0" rl="0" enabled="false"/>
    <control time="60" rl="5" enabled="true"/>
  </sensor>
  <sensor x="70" y="70" rl="1">
    <control time="0" rl="0" enabled="false"/>
    <control time="120" rl="7" enabled="true"/>
    <control time="180" rl="0" enabled="false"/>
  </sensor>
</network>
```

4 SIMULATING A WSN WITH MOBSINK

A simulation of a generic WSN deployed on a city will be presented, demonstrating the operation of the MobSink tool. In this section, a map of the city of Feira de Santana (Brazil) is considered, assuming that sensors will be deployed in this city. The MobSink can be used as shown in Figure 3 (main screen of the tool) to model every street of the city or the user can convert maps exported from OpenStreetMaps directly to the MobSink format, which is a feature also provided by the MobSink tool.

Once the user has created a map for simulation or imported a map from OpenStreetMaps, he/she can insert sensor nodes to create a WSN. There are three ways to insert sensor nodes:

- Manually: the user can click the sensor tool, type a relevance level in "RL" text box and click the position on the map where he/she wants to insert the sensor.
- Grid: the user can type the amount of sensors to be inserted into the "Sensors" text box and click the "Insert sensors in a grid manner" button. The sensors will be inserted onto the map forming a regular grid.
- Randomly: the user can type the amount of sensors to be inserted into the "Sensors" text box and click the "Insert sensors randomly" button. The sensors will be inserted onto the map in random positions.

In any of these options, there is also a possibility to randomly select the relevance level of the new sensor nodes by clicking the "Insert sensors with random RLs" button. In this case, every time the user inserts a new node, its relevance level will be randomly chosen.

After inserting sensors nodes, the sinks have to be inserted. The current version of MobSink does not allow to insert sinks in user-defined positions. Instead, the user has to type how many sinks he/she wants to insert into the "Sinks" text box, choose a sink positioning algorithm from the combo box labeled "Sink positioning" and click the "car" button. MobSink will place the sinks according to the chosen sink positioning method.

At last, the user can type the simulation time in seconds into the "Time" text box and start a simulation. To do so, the user has to click the "Start simulation" button. MobSink will simulate the network for the time the user has chosen and display its results in the text box at the right corner of the main window.

After a simulation had run, it is possible to save all the simulation data into a CSV file by clicking the "Save report" button. This file can be opened in most spreadsheet applications and can be parsed to generate graphics of energy consumption in the WSN. Along with energy and transmission data, it contains every movement the sinks may have done during the simulation. Figure 4 contains an example of a chart created with GNU Plot just processing MobSink provided data.

Also, it is possible to run an animated simulation if the user clicks the "Enable sinks animation" button. Sometimes, this is worth to watch the sinks moving during the simulation (the paths chosen by the sinks are highlighted every time they move). Actually, this is a helpful resource provided by our tool that can support different teaching methods in networks and WSN classes.

Finally, if the user needs to run several simulations in a batch manner, MobSink offers a command line interface.

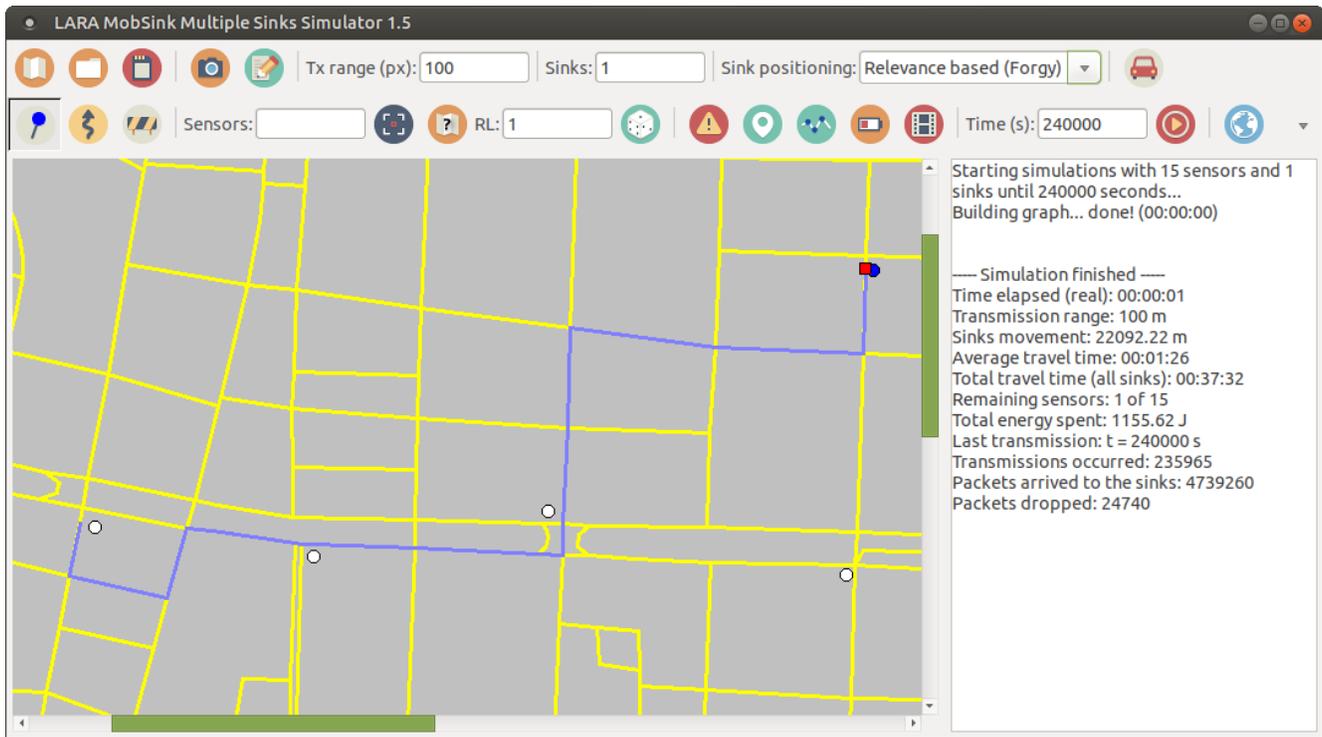


Figure 3: MobSink main screen.

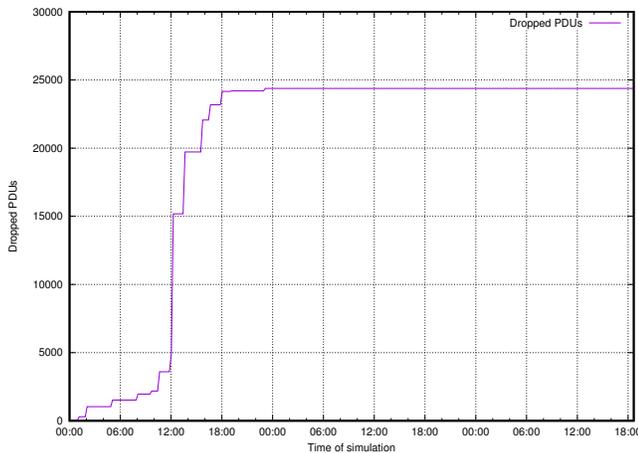


Figure 4: Number of dropped PDUs.

5 CONCLUSIONS

This paper presented MobSink, a novel tool to simulate multiple sinks positioning in WSNs and associated energy consumption. MobSink can be used to simulate dynamic sensors behavior and traffic jams, allowing the user to assess energy consumption and packet losses. The energy consumption accounting and sinks positioning algorithms were described in this paper, as well a flowchart explaining how the tool works when simulating a WSN. Also, the mechanism for sensors and traffic configurations were shown.

We believe MobSink can be very useful to students who want to better understand the basics of WSNs, sinks positioning and Smart Cities. And this will be even more evident in future versions of the tool, which will include additional features in this sense.

6 DOWNLOAD

The MobSink software can be downloaded from <http://just.pro.br/blog/mobsink>

REFERENCES

- [1] Y. Charfi, N. Wakamiya, and M. Murata. 2009. Challenging Issues in Visual Sensor Networks. *Wireless Commun.* 16, 2 (April 2009), 44–49.
- [2] E Egea-Lopez, J Vales-Alonso, AS Martinez-Sala, P Pavon-Marino, and J García-Haro. 2005. Simulation tools for wireless sensor networks. In *Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS)*. 24.
- [3] G. P. Hancke and G P. Silva, B. and Hancke Jr. 2012. The Role of Advanced Sensing in Smart Cities. *Sensors* 13, 1 (2012), 393.
- [4] M. Jevtic, N. Zogovic, and G. Dimic. 2009. Evaluation of Wireless Sensor Network Simulators. 17th Telecommunications Forum.
- [5] J. P. J. Peixoto and D. G. Costa. 2015. QoE-aware multiple sinks mobility in wireless sensor networks. In *Conference on New Technologies, Mobility and Security (NTMS)*. 1–4.
- [6] J. P. J. Peixoto and D. G. Costa. 2017. Wireless visual sensor networks for smart city applications: A relevance-based approach for multiple sinks mobility. *Future Generation Computer Systems* 76 (2017), 51 – 62.
- [7] D. Rosario, Z. Zhao, C. Silva, E. Cerqueira, and T. Braun. 2013. An OMNeT++ Framework to Evaluate Video Transmission in Mobile Wireless Multimedia Sensor Networks. In *Conference on Simulation Tools and Techniques (ICST)*.
- [8] H. Sundani, H. Li, V. Devabhaktuni, M. Alam, and P. Bhattacharya. 2008. Wireless Sensor Network Simulators: A Survey and Comparisons. 2 (April 2008). *International Journal Of Computer Networks*.
- [9] J. Yick, B. Mukherjee, and D. Ghosal. 2008. Wireless sensor network survey. *Computer Networks* 52, 12 (aug 2008), 2292–2330.

Integração de um Gerador de Código ao FrameWeb Editor

Breno Leite Zupeli

Núcleo de Estudos em Modelagem Conceitual e Ontologias (NEMO), Departamento de Informática, Universidade Federal do Espírito Santo (UFES) - Vitória, ES, Brasil
brenolzupeli@gmail.com

Vítor E. Silva Souza

Núcleo de Estudos em Modelagem Conceitual e Ontologias (NEMO), Departamento de Informática, Universidade Federal do Espírito Santo (UFES) - Vitória, ES, Brasil
vitor.souza@ufes.br

ABSTRACT

FrameWeb (Framework-based Design Method for Web Engineering) incorporates concepts from categories of frameworks commonly used in the development of Web-based Information Systems into design models, defining the syntax of such models with meta-models. Based on Model-Driven Development (MDD) techniques, a CASE tool called *FrameWeb Editor* was built. In a separate effort, a code generation tool was proposed, but did not use the method's MDD foundations. In this paper, we report on the integration of the code generator into the *FrameWeb Editor* and the *FrameWeb* meta-model.

KEYWORDS

Web Engineering, Frameworks, Model-Driven Development, *FrameWeb*, CASE Tool

1 INTRODUÇÃO

No contexto da Engenharia Web, é muito comum o reuso de código de infraestrutura por meio do uso de *frameworks* (ex.: Spring Framework¹) ou plataformas de desenvolvimento (ex.: Java EE²). Um *framework* é um artefato de código que provê componentes prontos que podem ser reutilizados mediante configuração, composição ou herança [18]. Seu uso auxilia a evitar a contínua redescoberta e reinvenção de padrões e componentes arquiteturais básicos, reduzindo custos e melhorando a qualidade do software, devido à reutilização de arquiteturas bem estabelecidas e testadas [17].

Com esta motivação, o método *FrameWeb* [18, 19] (*Framework-based Design Method for Web Engineering*) define uma arquitetura padrão para facilitar a integração com tais *frameworks*, além de propor um conjunto de modelos que traz para o projeto arquitetural do sistema conceitos inerentes a eles. O método é baseado em conceitos de MDD (*Model-Driven Development*) [16], sendo a sintaxe de sua linguagem de modelagem especificada formalmente por meio de meta-modelos [12], o que permite que suporte a novos *frameworks* sejam incluídos e os modelos de projeto sejam adaptados ao conjunto de *frameworks* que se deseja utilizar.

A abordagem MDD promove, também, a construção de ferramentas de apoio ao uso do método. Campos & Souza propõem o *FrameWeb Editor* [5], uma ferramenta CASE que dá suporte à criação e verificação dos modelos propostos por *FrameWeb* de forma gráfica e baseada nos meta-modelos que definem a sintaxe de sua

¹<https://spring.io>.

²<http://www.oracle.com/technetwork/java/javaee/overview/index.html>.

linguagem. O editor foi construído com base no projeto Sirius [20], que por sua vez é baseado no *Eclipse Modeling Framework* (EMF) [9], que permite o desenvolvimento de ferramentas dentro da abordagem MDD na plataforma Eclipse.³

Em um esforço paralelo, Almeida et al. [7] propõem um gerador capaz de criar esqueletos de código a partir dos modelos prescritos por *FrameWeb*, adaptados ao conjunto de *frameworks* escolhido para o projeto em questão. No entanto, o gerador de código não foi desenvolvido dentro da abordagem MDD, i.e., não utiliza o meta-modelo para especificar as extensões da linguagem necessárias para a geração de código (ex.: *templates* de código utilizados pelo gerador), nem faz uso de técnicas e ferramentas de transformação oferecidas pela abordagem MDD e pela plataforma Eclipse, tendo sido desenvolvido na plataforma .NET como uma ferramenta separada do *FrameWeb Editor*.

O objetivo deste artigo é apresentar os resultados de um esforço de integração do gerador de código [7] ao *FrameWeb Editor* [5], detalhando as modificações necessárias nos meta-modelos de *FrameWeb* para a adequação do gerador à abordagem MDD. A Seção 2 introduz o referencial teórico, a Seção 3 apresenta as mudanças feitas no *FrameWeb Editor* e nos meta-modelos de *FrameWeb* para integração do gerador de código, a Seção 4 compara trabalhos relacionados e a Seção 5 traz as considerações finais.

2 REFERENCIAL TEÓRICO

Para auxiliar na construção de Sistemas de Informação Web (*Web-based Information Systems* ou WIS) que possuam uma infraestrutura arquitetônica baseada no uso de *frameworks*, o método *FrameWeb* (*Framework-based Design Method for Web Engineering*) [18, 19] propõe uma arquitetura de software padrão, que divide o sistema em três grandes camadas — lógica de apresentação, lógica de negócio e lógica de acesso aos dados — seguindo o padrão arquitetônico *Service Layer* (Camada de Serviço) [8].

Com base nessa arquitetura, *FrameWeb* define quatro tipos de diagramas, todos baseados no Diagrama de Classes da UML [15], que são utilizados para representar os componentes específicos de cada camada em questão, abrangendo os elementos típicos da plataforma Web e relacionados aos *frameworks* utilizados.

Na camada de apresentação, o pacote View (visão) contém os elementos de interface gráfica com o usuário para a plataforma Web, como páginas HTML, imagens, folhas de estilo, scripts JavaScript, etc. O pacote Controller (controle) engloba as classes de controle, responsáveis por capturar os estímulos e os dados enviados pelo usuário por meio dos componentes de visão, chamar os serviços apropriados na aplicação e apresentar o resultado de volta ao usuário. Tal arquitetura é baseada nos *frameworks* Web

³<http://www.eclipse.org>.

(ou controlador frontal, ex.: *JavaServer Faces*⁴) e representada no **Modelo de Navegação**.

Na camada de negócio, o pacote Domain (domínio) contém as classes que representam os elementos do domínio do problema (conceitos do mundo real), cuja persistência é realizada por *frameworks* de mapeamento objeto/relacional (ex.: *Java Persistence API*⁵). Tais classes e seus mapeamentos são representados no **Modelo de Entidades**. Já o pacote Application (aplicação) reúne as classes de serviço, responsáveis pela implementação dos casos de uso (funcionalidades do WIS), que são representadas no **Modelo de Aplicação**, juntamente com suas dependências, gerenciadas pelo *framework* de injeção de dependências (ex.: *Contexts and Dependency Injection for Java*⁶).

Por fim, na camada de acesso a dados, o pacote Persistence (persistência) traz classes responsáveis pela persistência dos objetos de domínio, segundo o padrão de projeto *Data Access Object* (DAO) [1]. Os DAOs são representados no **Modelo de Persistência**.

Por serem baseados no Diagrama de Classes da UML, qualquer editor UML pode ser utilizado para construir os modelos descritos acima. No entanto, para garantir que os modelos sejam construídos utilizando apenas construtos válidos para o FrameWeb, a sintaxe da *linguagem FrameWeb* foi especificada formalmente por meio de meta-modelos [12], seguindo a abordagem MDD [16].

O Desenvolvimento Orientado a Modelos (*Model-Driven Development*, ou MDD) propõe uma abordagem na qual os modeladores concentram esforços na elaboração de modelos conceituais para representação de todas as características de um sistema, e possibilita modelar as funcionalidades e recursos desejáveis para o sistema, portanto, o modelo gerado é o projeto do sistema em si. Posteriormente, por meio de transformações entre modelos de diversos níveis de abstração é feita a implementação do sistema. Algumas das principais vantagens do MDD para o processo de geração de código são: **portabilidade**, porque os modelos de alto nível podem ser transformados em códigos diferentes, dependendo da plataforma ou *framework* escolhidos; **corretude**, pois geradores de código não produzem erros acidentais e além disso facilitam a identificação de erros conceituais, visto que ocorrem em níveis mais altos de abstração; e **otimização**, dado que os modelos possibilitam mais eficiência e menor incidência de erros aos geradores.

A partir do meta-modelo de FrameWeb, foi construído o *FrameWeb Editor*, que oferece um ambiente gráfico que possibilita a elaboração destes diagramas, bem como a definição da plataforma de programação e dos *frameworks* utilizados. A Figura 1 apresenta uma visão geral do editor, que possui do lado esquerdo o *Project Explorer* para visualização do projeto, no lado direito apresenta um painel com as opções para criação de componentes de acordo com o modelo utilizado, e na parte de baixo são mostradas as opções para configuração do atributo selecionado. Um duplo-clique em um dos componentes de modelo permite ao desenvolvedor construir o modelo em questão. A figura mostra, também, que para este projeto foi escolhida a plataforma Java e o *framework* JSF.

A fim de facilitar o desenvolvimento de projetos elaborados no *FrameWeb Editor*, um gerador de código foi desenvolvido para geração do esqueleto do código, baseando-se no arquivo gerado

⁴JSF, <http://jcp.org/en/jsr/detail?id=344>

⁵JPA, <http://jcp.org/en/jsr/detail?id=338>

⁶CDI, <http://jcp.org/en/jsr/detail?id=346>

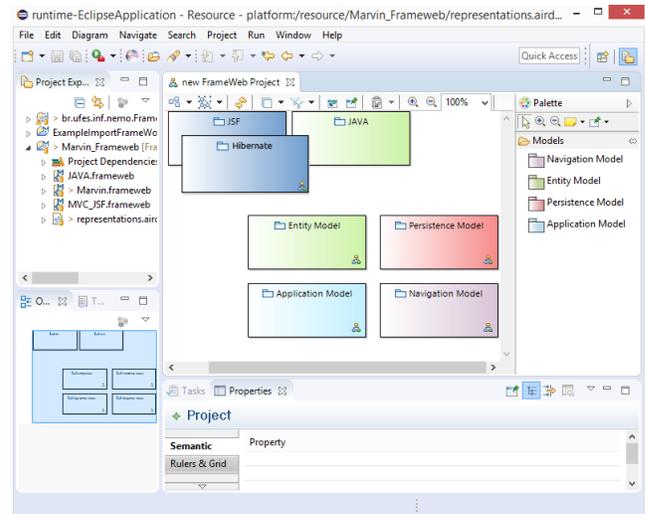


Figura 1: Visão geral do FrameWeb Editor [5].

pelo editor com as informações dos modelos criados, visto que este gerador não se encontrava integrado ao *FrameWeb Editor*. A próxima sessão apresenta como foi feita essa integração.

3 INTEGRAÇÃO DO GERADOR DE CÓDIGO

Partiu-se do *FrameWeb Editor* [5] e do gerador de código [7], reimplementando este último usando a linguagem Java (sua implementação original foi feita em C#), para possibilitar sua integração ao meta-modelo do FrameWeb e, assim, ao editor.

O gerador de código em sua versão original armazena os *templates* por meio de arquivos organizados em diretórios, um arquivo de configuração deve ser modificado para indicar os diretórios dos *templates* necessários e o diretório onde o código gerado será armazenado, isso não permite uma boa usabilidade ao usuário, além de não promover uma ligação explícita destes *templates* com a linguagem FrameWeb, que é definida por meta-modelos.

A fim de permitir a utilização de vários tipos de *frameworks* e de integrar os *templates* ao meta-modelo do FrameWeb, foram adicionadas três novas meta-classes, representadas na Figura 2: **FrontControllerTemplate**, responsável por armazenar *templates* das páginas, dos componentes das páginas, das classes controladoras, seus métodos e atributos; **DI Template**, responsável por armazenar *templates* das classes de serviço, seus métodos e atributos; e **ORMTemplate**, responsável por armazenar *templates* das classes de domínio e seus atributos. Além disso, a classe **Tag**, já existente, ganhou o atributo `codeGenerationTemplate` para armazenamento dos *templates* das *tags* HTML utilizadas nos Modelos de Navegação.

Instâncias destas novas meta-classes podem ser vistas na Figura 1, apresentada anteriormente. O elemento JSF representa uma instância da classe `FrontControllerTemplate` e armazena os *templates* das classes, métodos, páginas, etc. Já Hibernate é instância da classe `ORMTemplate` e armazena os *templates* de classes e atributos. Estes *frameworks* são utilizados nos modelos de Navegação e Entidades, respectivamente.

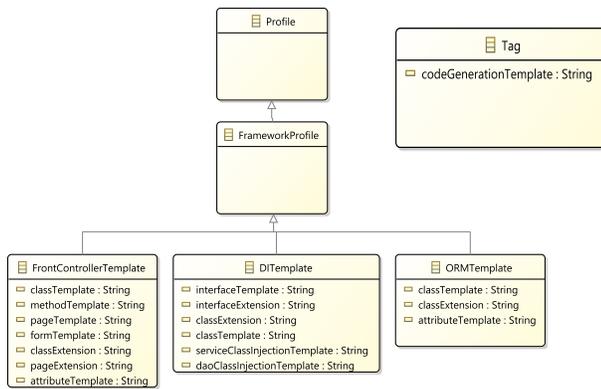


Figura 2: Alterações feitas no meta-modelo de FrameWeb.

Devido ao fato de possuírem caracteres especiais, tais *templates* são transformados por meio de um codificador de URLs (já disponível na API da linguagem Java) para armazenamento nos atributos das respectivas meta-classes, sendo decodificados em tempo de execução pelo gerador de código. Os *templates* utilizam variáveis que possuem o prefixo FW_, cada uma delas correspondendo a um atributo de uma meta-classe do meta-modelo de FrameWeb, ou seja, referente a alguma informação que será extraída dos modelos criados com o FrameWeb Editor, conforme já apresentado em [7].

A Figura 3 mostra um exemplo de Modelo de Navegação, representando uma funcionalidade do Sistema de Controle de Afastamento de Professores (SCAP), uma aplicação que auxilia departamentos de universidades federais a gerenciar pedidos de afastamento (*leave of absence requests*) de seus professores. O modelo traz uma página chamada form.xhtml, que possui um formulário a ser preenchido pelo professor que deseja solicitar seu afastamento, cujos dados são enviados à classe controladora RequestLeaveController, que tem o seu método submit() associado a esta submissão. Ao final do registro da solicitação de afastamento, o controlador redireciona o usuário para a página success.xhtml.

Com a integração do gerador de código ao FrameWeb Editor, para ativar a geração de código, o usuário deve, na visão geral do projeto no editor (Figura 1), clicar com o botão direito e selecionar a opção *Gerar código*. Uma janela é exibida para a seleção do diretório desejado para a criação do projeto com o código gerado. Na sequência, detalhamos como esta geração de código acontece utilizando o exemplo da Figura 3.

Para que as páginas HTML e seus formulários («page» e «form» no Modelo de Navegação) sejam gerados a partir do modelo, na definição do *framework* JSF (Figura 1), como instância da meta-classe FrontControllerTemplate (Figura 2), os atributos pageTemplate e formTemplate foram preenchidos, respectivamente, com os conteúdos das listagens 1 e 2. Na página Web, a variável FW_BODY é substituída por formulários que compõem a página, enquanto que esta mesma variável no *template* do formulário é substituída pelos seus componentes visuais (representados por seus atributos no Modelo de Navegação).

Também na definição do JSF, as diferentes *tags* que representam os componentes visuais utilizados em formulários Web providas

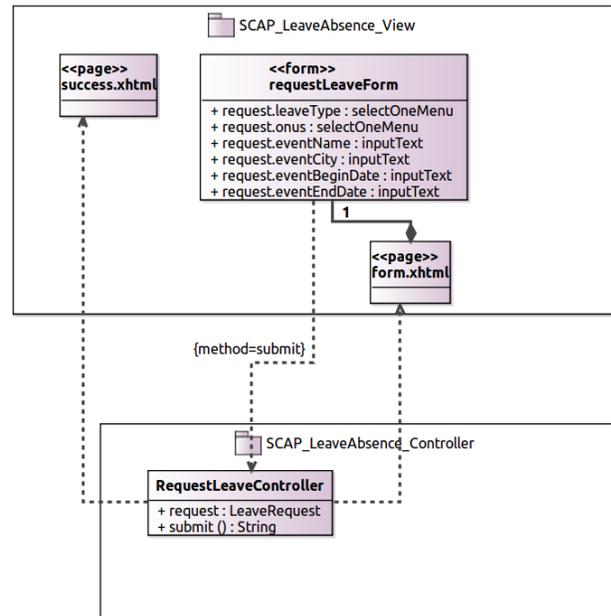


Figura 3: Exemplo de um modelo de Navegação.

Listagem 1: *Template* para páginas HTML (pageTemplate).

```

<?xml version="1.0" encoding="ISO-8859-1" ?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.
w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml"
xmlns:f="http://java.sun.com/jsf/core"
xmlns:h="http://java.sun.com/jsf/html"
xmlns:ui="http://java.sun.com/jsf/facelets">
<head>
<meta http-equiv="Content-Type" content="text/html; charset=UTF-8" />
<title>Insert title here</title>
</head>
<h:body>
FW_BODY
</h:body>
</html>
    
```

Listagem 2: *Template* para form HTML (formTemplate).

```

<h:form id="FW_ID">
FW_BODY
</h:form>
    
```

Listagem 3: *Template* para campo JSF inputText.

```

<h:inputText id="FW_ID" value="#{FW_VALUE}" /><br />
    
```

pelos *framework* são registradas como instâncias da meta-classe Tag e também associadas a um *template* de código. A Listagem 3 mostra o *template* para componentes de texto simples (<h:inputText />) enquanto a Listagem 4 se refere a componentes de caixa de seleção (<h:selectOneMenu />). Neste caso, a variável FW_VALUE é substituída pelo nome do componente (atributo da classe «form» no Modelo de Navegação) para que o JSF faça a ligação (*binding*) entre o componente Web e o controlador Java.

Listagem 4: Template para campo JSF selectOneMenu.

```
<h:selectOneMenu id="FW_ID" value="#{FW_VALUE}"></h:selectOneMenu><
  ↪ br />
```

Listagem 5: Parte do arquivo form.xhtml gerado.

```
...
<h:body>
<h:form id="requestLeaveForm">
<h:selectOneMenu id="request_leaveType" value="#{requestLeaveForm.
  ↪ request.leaveType}"></h:selectOneMenu><br />
...
<h:inputText id="request_eventEndDate" value="#{requestLeaveForm.
  ↪ request.eventEndDate}" /><br />
</h:form>
</h:body>
</html>
```

Listagem 6: Template para controladora (classTemplate).

```
package FW_FRONT_CONTROLLER_PACKAGE;

import javax.faces.bean.ManagedBean;
import javax.faces.bean.SessionScoped;

@ManagedBean(name = "FW_BEAN_NAME") @SessionScoped
public class FW_CLASS_NAME {
    FW_FRONT_CONTROLLER_ATTRIBUTES
    public FW_CLASS_NAME() { }
    FW_FRONT_CONTROLLER_METHODS
}
```

Listagem 7: Template para atributo (attributeTemplate).

```
private FW_ATTRIBUTE_TYPE FW_ATTRIBUTE;
public FW_ATTRIBUTE_TYPE getFW_ATTRIBUTE_FIRST_UPPER() {
    return FW_ATTRIBUTE;
}
public void setFW_ATTRIBUTE_FIRST_UPPER(FW_ATTRIBUTE_TYPE
  ↪ _FW_ATTRIBUTE) {
    FW_ATTRIBUTE = _FW_ATTRIBUTE;
}
```

Ao utilizar todos estes *templates* para gerar código do elemento form.xhtml do modelo da Figura 3, parte do resultado é representada na Listagem 5.

Analogamente, para que a classe controladora seja gerada, a definição do *framework* JSF traz nos atributos classTemplate, attributeTemplate e methodTemplate, respectivamente, o conteúdo das listagens 6, 7 e 8. O *template* da classe traz as anotações do JSF para que a classe possa ser referenciada a partir das páginas Web, bem como variáveis que são substituídas pelo código gerado a partir dos atributos (FW_FRONT_CONTROLLER_ATTRIBUTES) e métodos (FW_FRONT_CONTROLLER_METHODS).

O *template* de atributos gera o código do atributo em si, juntamente com métodos de acesso (get/set) utilizados pelo JSF. Por fim, o *template* de métodos apenas inclui um esqueleto do método, sem implementação. Combinando tais *templates* na geração da classe RequestLeaveController da Figura 3, o resultado é o exibido pela Listagem 9.

Conforme proposto em [12], o FrameWeb pode ser estendido por meio da criação de novos perfis de *framework*. Com a integração do gerador de código ao FrameWeb Editor, tal esforço envolve, agora, instanciar as classes da Figura 2 e prover os *templates* de código para que o gerador possa construir, a partir dos respectivos modelos,

Listagem 8: Template para método (methodTemplate).

```
public FW_METHOD_RETURN_TYPE FW_METHOD_NAME(){
    return null;
}
```

Listagem 9: Classe RequestLeaveController gerada.

```
package br.ufes.inf.scap.control;

import javax.faces.bean.ManagedBean;
import javax.faces.bean.SessionScoped;

@ManagedBean(name = "requestLeaveController") @SessionScoped
public class RequestLeaveController {
    private LeaveRequest request;
    public LeaveRequest getRequest() {
        return request;
    }
    public void setRequest(LeaveRequest _request){
        request = _request;
    }
    public RequestLeaveController() { }
    public String submit() {
        return null;
    }
}
```

os artefatos de código de maneira adequada, conforme ilustrado acima para o *framework* JSF. O próprio editor provê suporte para definição de novos *frameworks* em sua interface gráfica [5].

A ferramenta FrameWeb Editor, que possui o gerador de código integrado, encontra-se sob a licença MIT e o leitor interessado pode obter instruções detalhadas de instalação e uso na página do repositório do projeto, hospedada em <https://github.com/nemo-ufes/FrameWeb/wiki/A-Ferramenta-FrameWeb-Editor>.

4 TRABALHOS RELACIONADOS

Assim como FrameWeb, outros métodos de Engenharia Web possuem ferramentas CASE e/ou geradores de código associados.

O método UWE (UML-based Web Engineering) [10] que, assim como FrameWeb, é baseado em MDD, possui a ferramenta CASE MagicUWE,⁷ porém a mesma não oferece suporte a geração de código. Uma outra ferramenta, UWE4JSF,⁸ permite gerar aplicações JSF baseada numa versão estendida do perfil UWE, porém não é integrada ao editor. Outro método MDD, o OOH4RIA [13], focado no desenvolvimento de *Rich Internet Applications*, possui a ferramenta associada OIDE [14] (OOH4RIA IDE), que gera código para os *frameworks* Google Web Toolkit e Silverlight.

O RUX-Method [4], focado em criação de interfaces sensíveis a contexto, possui uma ferramenta chamada RUX-Tool [11], capaz de gerar código de componentes de interface com o usuário automaticamente, devendo porém ser usada em conjunto com a ferramenta comercial WebRatio.⁹ Esta ferramenta utiliza a linguagem visual IFML [2] para modelagem de fluxos de interação, controle de comportamento e conteúdo de interfaces com o usuário (*front-ends*), linguagem adotada pela OMG¹⁰ como padrão em 2013. Alternativas de código aberto para edição deste tipo de modelo, como o IFMLEdit.org [3] também oferecem geração de código.

⁷<http://uwe.pst.ifi.lmu.de/toolMagicUWE.html>.

⁸<http://uwe.pst.ifi.lmu.de/toolUWE4JSF.html>.

⁹<https://www.webratio.com>.

¹⁰<https://www.omg.org/>.

Há geradores de código que trabalham com modelos UML e arquivos de mapeamento (por exemplo, em XML) para a geração de código de aplicações Web. GenERTiCA [21] se propõe a gerar código de forma mais completa possível para sistemas de tempo real distribuídos, utilizando conceitos da Programação Orientada a Aspectos. WebML [6] foca na especificação visual de websites complexos, permitindo a geração de código das páginas HTML que compõem seu *front-end*.

O principal diferencial da abordagem FrameWeb — incluindo seu editor e gerador de código, integrados neste trabalho — em relação a outras abordagens, inclusive ferramentas não-acadêmicas como, por exemplo, JHipster,¹¹ é a sua extensibilidade. Seus metamodelos permitem que os usuários do FrameWeb Editor adicionem suporte a novos *frameworks* (inclusive legados, que dificilmente teriam suporte de ferramenta), incluindo *templates* que possibilitam a geração de código para tais *frameworks*. Tais *templates* também podem ser personalizados pelos usuários do FrameWeb Editor.

5 CONCLUSÕES

Neste artigo, apresentamos o resultado de um esforço de integração entre a ferramenta CASE FrameWeb Editor [5] e o gerador de código [7] proposto para o método FrameWeb, que se fez necessário pois este último não foi construído com base no meta-modelo FrameWeb [12].

A integração se faz necessária para que os usuários do método tenham uma interface unificada para aproveitar a principal vantagem de FrameWeb, sua extensibilidade: os próprios usuários finais do FrameWeb Editor podem cadastrar novos *frameworks* e modificar os *templates* para geração de código, adaptando-os às necessidades do projeto ou da organização na qual trabalham. Em particular, projetos legados que utilizam *frameworks* que não possuem suporte de ferramentas comerciais ou *open source* disponíveis podem ser beneficiar desta capacidade de FrameWeb de se adaptar a diferentes conjuntos de *frameworks*.

A integração do editor e do gerador em uma única ferramenta traz como vantagens uma maior facilidade de uso, pois o usuário final não precisa instalar e usar duas ferramentas separadas, de plataformas diferentes; diminuição da ambiguidade em modelos, pois são gerenciados por uma única ferramenta; e uma manutenção mais fácil para os desenvolvedores FrameWeb, que não precisam manter duas bases de código diferentes, em linguagens de programação diferentes. A construção do gerador dentro da plataforma Eclipse permitirá, também, o uso dos recursos de Desenvolvimento Orientado a Modelos presentes na plataforma.

Como trabalhos futuros, pretende-se utilizar o FrameWeb Editor e seu gerador de código integrados na atual edição do curso de *Desenvolvimento Web e Web Semântica*¹² em nossa universidade, avaliando-o junto aos alunos matriculados nesta disciplina. Também encontra-se em andamento, no contexto de um projeto de Iniciação Científica, uma maior integração do FrameWeb Editor com a IDE Eclipse, de modo a aproximar a ferramenta da indústria de desenvolvimento de software, promovendo transferência de tecnologia da pesquisa sobre FrameWeb da academia para as empresas do ramo e permitindo formas mais robustas de avaliação da mesma.

¹¹<https://jhipster.tech>.

¹²<https://www.inf.ufes.br/~vitorsouza/teaching/dwvs-20182/>.

ACKNOWLEDGMENTS

O NEMO (<http://nemo.inf.ufes.br>) possui atualmente apoio do CNPq (processo 407235/2017-5), da CAPES (23038.028816/2016-41) e da FAPES (69382549/2015). Este trabalho foi realizado com bolsa do Programa Institucional de Iniciação Científica da UFES.

REFERÊNCIAS

- [1] Deepak Alur, John Crupi, and Dan Malks. 2003. *Core J2EE Patterns: Best Practices and Design Strategies* (2nd ed.). Prentice Hall / Sun Microsystems Press.
- [2] Luciano Baresi, Franca Garzotto, and Paolo Paolini. 2001. Extending UML for modeling web applications. In *System Sciences, 2001. Proceedings of the 34th Annual Hawaii International Conference on*. IEEE, 10–pp.
- [3] Carlo Bernaschina, Sara Comai, and Piero Fraternali. 2017. IFMLEdit.Org: Model Driven Rapid Prototyping of Mobile Apps. In *Proceedings of the 4th International Conference on Mobile Software Engineering and Systems (MOBILESoft '17)*. IEEE Press, 207–208. <https://doi.org/10.1109/MOBILESoft.2017.15>
- [4] Gaëlle Calvary, Joëlle Coutaz, David Thevenin, Quentin Limbourg, Laurent Bouillon, and Jean Vanderdonckt. 2003. A unifying reference framework for multi-target user interfaces. *Interacting with computers* 15, 3 (2003), 289–308.
- [5] Silas Louzada Campos and Vitor E. S. Souza. 2017. FrameWeb Editor: Uma Ferramenta CASE para suporte ao Método FrameWeb. In *Anais do 16º Workshop de Ferramentas e Aplicações, 23º Simpósio Brasileiro de Sistemas Multimedia e Web*. SBC, Gramado, RS, Brazil, 199–203.
- [6] Stefano Ceri, Piero Fraternali, and Aldo Bongio. 2000. Web Modeling Language (WebML): a modeling language for designing Web sites. *Computer Networks* 33, 1–6 (jun 2000), 137–157. [https://doi.org/10.1016/S1389-1286\(00\)00040-2](https://doi.org/10.1016/S1389-1286(00)00040-2)
- [7] Nilber V. de Almeida, Silas L. Campos, and Vitor E. S. Souza. 2017. A Model-Driven Approach for Code Generation for Web-based Information Systems Built with Frameworks. In *Proc. of the 23rd Brazilian Symposium on Multimedia and the Web*. ACM, Gramado, RS, Brazil, 245–252. <https://doi.org/10.1145/3126858.3126863>
- [8] Martin Fowler. 2002. *Patterns of Enterprise Application Architecture* (1 ed.). Addison-Wesley.
- [9] Richard C. Gronback. 2009. *Eclipse Modeling Project: A Domain-Specific Language (DSL) Toolkit* (1 ed.). Addison-Wesley Professional.
- [10] Nora Koch, Alexander Knapp, Gefei Zhang, and Hubert Baumeister. 2008. Web Engineering: Modelling and Implementing Web Applications. In *UML-Based Web Engineering*. Springer, London, UK.
- [11] Marino Linaje, Juan Carlos Preciado, Rober Morales-Chaparro, Roberto Rodriguez-Echeverria, and Fernando Sánchez-Figueroa. 2009. Automatic Generation of RIAs Using RUX-Tool and Webratio. In *Web Engineering: 9th International Conference, ICWE 2009, Proceedings*. Springer Berlin Heidelberg, San Sebastián, Spain, 501–504. https://doi.org/10.1007/978-3-642-02818-2_48
- [12] Beatriz Franco Martins and Vitor E. S. Souza. 2015. A Model-Driven Approach for the Design of Web Information Systems based on Frameworks. In *Proc. of the 21st Brazilian Symposium on Multimedia and the Web*. ACM, 41–48.
- [13] Santiago Meliá, Jaime Gómez, Sandy Pérez, and Oscar Diaz. 2008. A model-driven development for GWT-based rich internet applications with OOH4RIA. In *Web Engineering, 2008. ICWE'08. Eighth International Conference on*. IEEE, 13–23.
- [14] Santiago Meliá, Jose-Javier Martínez, Sergio Mira, Juan Antonio Osuna, and Jaime Gómez. 2010. An Eclipse Plug-in for Model-Driven Development of Rich Internet Applications. In *Web Engineering*. Springer, 514–517.
- [15] OMG. 2017. Unified Modeling Language, version 2.5.1. <https://www.omg.org/spec/UML/2.5.1/>. (2017).
- [16] Oscar Pastor, Sergio España, José Ignacio Panach, and Nathalie Aquino. 2008. Model-driven development. *Informatik-Spektrum* 31 (2008), 394–407.
- [17] D.C. Schmidt, M. Stal, H. Rohnert, and F. Buschmann. 2013. *Pattern-Oriented Software Architecture, Patterns for Concurrent and Networked Objects*. Wiley.
- [18] Vitor E Silva Souza. 2007. *FrameWeb: um Método baseado em Frameworks para o Projeto de Sistemas de Informação Web*. Master's thesis. Programa de Pós-Graduação em Informática, Universidade Federal do Espírito Santo.
- [19] Vitor E. S. Souza, Ricardo A. Falbo, and Giancarlo Guizzardi. 2009. Designing Web Information Systems for a Framework-based Construction. In *Innovations in Information Systems Modeling: Methods and Best Practices* (1 ed.), Terry Halpin, Eric Proper, and John Krogstie (Eds.). IGI Global, Chapter 11, 203–237.
- [20] Vladimir Vojovic, Milan Maksimovic, and Branko Perisic. 2014. Sirius: A rapid development of DSM graphical editor. In *Intelligent Engineering Systems (INES), 2014 18th International Conference on*. IEEE, 233–238.
- [21] Marco A. Wehrmeister, Edison P. Freitas, Carlos E. Pereira, and Franz Rammig. 2008. GenERTiCA: A Tool for Code Generation and Aspects Weaving. In *Proc. of the 11th IEEE Symposium on Object Oriented Real-Time Distributed Computing (ISORC)*.

VolleyJump: Uma aplicação para a análise de saltos no voleibol de praia

Renan Bandeira,
Fernando Trinta,
João Gomes
Universidade Federal do Ceará
Av. Humberto Monte, SN - PICI
Fortaleza, Ceará
renanbandeira@alu.ufc.br,fernando.
trinta@dc.ufc.br,jpaulo@lia.ufc.br

Marcio Maia
UFC - Campus Quixadá
Av. José de Freitas Queiroz, 5003
Quixadá, Ceará
marcioefmaia@gmail.com

Alexandre Araripe
Universidade Federal do Ceará
Av. Humberto Monte, SN - PICI
Fortaleza, Ceará
alexandremedeiros@ufc.br

ABSTRACT

Professional sports are increasingly dependents of technological resources given the remarkable level of competitiveness faced by high performance athletes. With such resources, it is possible to analyze matches, avoid mistakes that may be committed by the referee or even to analyze the athletes' performance. One of these sports is beach volleyball, one of most popular sports in Brazil. In the past 12 years, the Brazilian volleyball teams has been always among the best teams in the world. The athletes' performance during the jump movement is one of the main factors that one team needs to improve to be successful because it is the movement that is most performed during a volleyball match. There are some approaches that study the jump movement in order to calculate its height and give evidences to improve it. Nevertheless, these solutions are expensive and are not viable to athletes with no sponsorship. Having this in mind, this works presents VolleyJump, an application created to analyze beach volleyball athlete jumps using machine learning strategies to calculate the jump height and classify it as an attack or block jump. Results show that VolleyIoT makes possible to analyze athletes' jumps using mobile devices sensors, helping them to focus on their training to improve its technique.

KEYWORDS

Wearable Computing, IoT, Mobile Device

1 INTRODUÇÃO

Atualmente, o esporte de alto rendimento está cada vez mais dependente de recursos tecnológicos. É difícil pensar em algum esporte em que um aparato tecnológico não esteja por trás de recordes e melhorias de desempenho dos atletas, ou mesmo para que uma comissão julgadora possa avaliar a quebra de recorde, ou analisar os lances durante o jogo e não falhar nos julgamentos realizados. Um destes esportes é o voleibol de praia.

No Voleibol de praia, boa parte das ações básicas de um jogador envolvem saltos. Um atleta salta tanto para ações de ataque à quadra do adversário, quanto na tentativa de bloquear o ataque do oponente.

Um atleta de alto nível neste esporte precisa manter seu nível de saltos de modo a manter seu desempenho homogêneo durante todo o andamento de uma partida. Com isso, analisar os saltos pode gerar resultados muito importantes para a melhoria do desempenho de um atleta. Dentre algumas questões que podem ser obtidas com esta análise estariam: (i) Quantos saltos um atleta dá, em média, durante uma jogo ou treino? (ii) Qual a diferença entre um salto de ataque um salto de bloqueio? (iii) Como se mantém o desempenho dos saltos do atleta durante o jogo?

A resposta a tais perguntas é essencial para a análise de fadiga do atleta, além de facilitar um treino focado na melhoria de determinados movimentos, e consequente melhoria do desempenho de atletas de voleibol. Nos últimos anos, dispositivos vestíveis tem encontrado um novo nicho no mercado para o monitoramento de performance em esportes como o golfe, futebol, basquete, dentre outros. A partir de dados coletados de sensores em objetos como pulseiras e relógios inteligentes, é possível utilizar técnicas de aprendizado de máquina e técnicas supervisionadas de classificação, para identificar novos dados de saltos a partir de exemplos gerados previamente.

É neste contexto que este documento apresenta a VolleyJump, uma aplicação que se propõe a viabilizar a melhoria de rendimento dos atletas de voleibol de praia a partir do uso de técnicas de aprendizagem de máquina para analisar um treinamento ou jogo e direcionar o treinamento do atleta em deficiências encontradas, de acordo com os resultados da aplicação. Para isso, a aplicação busca coletar dados contextuais do atleta durante o movimento do salto e, a partir destas informações, gerar informações relevantes referentes ao movimento. Com estes dados, o VolleyJump tem como objetivo (a) inferir o momento em que um salto foi realizado, (b) calcular a altura do salto e gerar um modelo de aprendizagem de máquina para que, quando novos dados contextuais de um salto forem gerados, ser capaz de (c) classificá-lo entre salto de ataque ou bloqueio.

2 TRABALHOS RELACIONADOS

Identificar a ação de uma pessoa a partir de dados coletados por sensores e dispositivos atrelados ao corpo de um usuário não é algo novo. O tema tem ganhado cada vez mais destaque na área do esporte. [2] e [6] buscam analisar a performance e saúde de atletas durante a prática do futebol e *mountain bike*, respectivamente. [2] propõe um modelo de arquitetura onde é possível monitorar as condições de saúde do atleta para que seja possível evitar lesões

In: XVII Workshop de Ferramentas e Aplicações (WEA 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pôsteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.
© 2018 SBC – Sociedade Brasileira de Computação.
ISBN 978-85-7669-435-9.

futuras. Porém, [6], apresenta um sistema de sensores distribuído a fim de monitorar a performance do atleta de *mountain bike* com o objetivo de futuras melhorias. No voleibol de quadra, [3] busca identificar saltos realizados durante um jogo analisando a frequência dos dados do acelerômetro.

A análise de saltos no voleibol de quadra é o tema estudado em [7]. Neste trabalho, os autores buscam encontrar alternativas à identificação de um salto de voleibol e o cálculo da altura de maneira rápida, de modo que o atleta possa realizar movimentos durante a prática do esporte e o treinador possa estar ciente do seu desempenho no momento da partida. Para isso, foi desenvolvido um dispositivo vestível que captura os dados do acelerômetro no eixo perpendicular à superfície para que pudesse ser calculado o que o trabalho propõe. Para que o treinador acompanhe o atleta, estes dados gerados são enviados via comunicação de rede sem fio (*wireless*) ao *smart phone*, possibilitando que o treinador já tenha a informação gerada a partir do dado segundos depois da realização do salto. A proposta aqui apresentada é semelhante ao proposto por [7], porém com a possibilidade também de se distinguir o tipo do salto do atleta, entre os de bloqueio e de ataque. Esta classificação automática pode ser de grande valia para o treinamento e melhoria do desempenho de um atleta, permitindo que fisiologistas e preparadores físicos possam analisar diferenças acentuadas em relação aos tipos de salto, ou mesmo uma queda de desempenho maior para um tipo de salto em relação a outro durante uma partida. Pelo que foi levantando na revisão de literatura relacionada ao tema, não existem soluções focadas em identificar o tipo de salto, um dos objetivos deste trabalho.

Com o foco em identificar padrões de estratégias bem sucedidas, [8] aplica dados da partida final da copa do mundo de voleibol de quadra de 2014 em técnicas de aprendizagem de máquina a fim de identificar padrões de ataques e defesas bem sucedidas durante o jogo. As estratégias encontradas são relevantes para que a equipe possa vencer o jogo, mas não são direcionadas para o rendimento e recuperação do atleta.

No voleibol de praia, o tema também é estudado. [4] busca criar uma rede neural para identificar atividades realizadas durante a prática do esporte, entre elas os saltos de levantamento e ações relacionadas ao movimento do braço, como o movimento do bloqueio e outros movimentos relacionados ao ataque, com uma acurácia de 83.2%. No entanto, a altura e os saltos de bloqueio não são levados em consideração.

Entre as técnicas utilizadas na academia, existem diversas delas que utilizam aparatos tecnológicos, em que é possível analisar os saltos de atletas de voleibol. Para medir a altura do salto, existe um dispositivo chamado *Vertec*, que o atleta salta e deve tocar em parte do aparato, como mostra a figura 1.

Além disso, existem plataformas de força que realizam a análise do salto baseado no tempo de voo do atleta, partindo do momento em que ele já não está mais pisando na plataforma até o momento onde ele volta a encostar nela [5]. A plataforma *Jump System Optical* usa sensores de presença para identificar os momentos de início do salto e queda em solo e a plataforma *Jump System Pro* usa sensores de pressão em sua plataforma para identificar o contato do atleta com o solo e, a partir do tempo de voo, calcular a altura do salto. Ambas plataformas dependem de um solo plano para a



Figura 1: Dispositivo Vertec responsável por medir a altura do atleta. Fonte: roguefitness.com

execução dos saltos, além de que passadas anteriores ao movimento devem ser executados fora da plataforma. Essas abordagens geram um resultado eficiente, mas são soluções caras para o problema proposto.

3 PROPOSTA

Para cada salto realizado, esta aplicação propõe identificá-los e mostrar resultados a partir da análise do salto executado. Para isso, o salto deve gerar dados, que devem ser analisados a fim de inferir a altura e o tipo de salto realizado. Portanto, é necessário dois passos para validar esta aplicação: (i) coleta e (ii) análise dos dados. Para o primeiro passo, a aplicação desenvolvida para a plataforma *Android* chamada *VolleyJump* permite a coleta de dados de sensores portados por um atleta, a fim de inferir dados sobre saltos. Os dados são enviados para um serviço de banco de dados na nuvem, responsável por armazenar os dados de maneira em que seja possível acessá-los a partir de qualquer dispositivo com acesso a Internet a qualquer momento. Com esses dados, um servidor na nuvem busca aplicar técnicas de aprendizagem de máquina e identificar o tipo e altura dos saltos, enviando essa informação ao dispositivo. A figura 2 apresenta este fluxo de troca de dados.

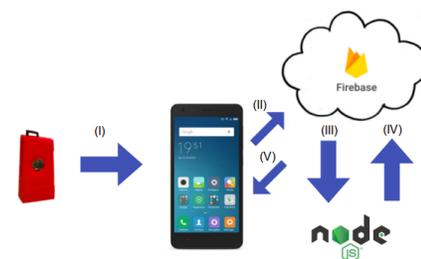


Figura 2: Fluxo de comunicação e troca de dados no Volley-Jump

No desenvolvimento da *VolleyJump*, os dados são coletados pelo sensor e repassados ao *smarthphone* (passo 1) e enviados a um banco de dados na nuvem, o *Firebase* (passo II). O servidor é notificado (passo III) e analisa os dados, retornando o resultado da análise

para Firebase (Passo IV). Por fim, a aplicação móvel é notificado ao final da análise (Passo V).

No caso desta aplicação, o diferencial em relação aos trabalhos relacionados é a categorização dos saltos em ataque ou bloqueio, utilizando assim técnicas de aprendizagem automática supervisionado de classificação. Este trabalho propõe o uso de uma estratégia de aprendizagem de máquina supervisionado, pois durante a execução do salto, é possível identificar o tipo e enviar ao algoritmo para que ele aprenda a partir de exemplos. Inicialmente, estes exemplos devem vir com o tipo e altura do salto, além dos dados gerados pelos sensores. Estes exemplos são essenciais para que a aplicação possa gerar um modelo de aprendizagem de máquina e, a partir de novos saltos, a aplicação possa inferir o tipo e a altura do salto. Para gerar estes exemplos, um grupo de possíveis atletas realizou saltos em uma plataforma usando o dispositivo que captura dados sensoriais a fim de que todos os dados necessários para a geração do modelo de aprendizagem de máquina usando o algoritmo de máquinas de vetor de suporte. Após gerar o modelo, novos saltos foram realizados sem o auxílio da plataforma de força para que a própria aplicação identifique e calcule a altura e o tipo de salto. Finalmente, os resultados são comparados.

4 VISÃO GERAL

Para que a aplicação possa contemplar o que este trabalho propõe, foi necessária a criação de duas telas: a primeira para a coleta de dados e a segunda para mostrar os resultados obtidos, como mostra a Figura 3.

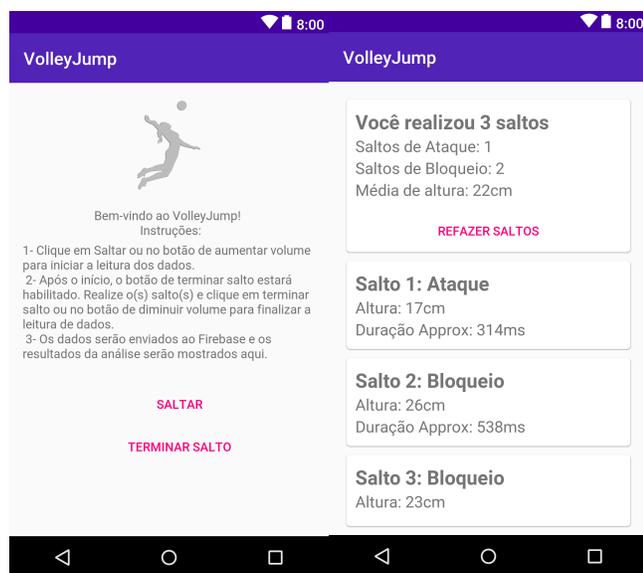


Figura 3: VolleyJump: tela responsável pela coleta de dados no lado esquerdo e os resultados no lado direito

4.1 Coleta de Dados

A fim de deixar o atleta confortável durante a prática do esporte, um dispositivo vestível é acoplado ao corpo do atleta e se comunica com a aplicação *Android* a fim de enviar os dados coletados. Essa

comunicação usa o padrão de projetos *Observer* [1] para o envio de dados gerados por sensores e, para isso, é preciso que o VolleyJump registre um *SensorTagListener*, que é uma interface que contém o método *onSensorTagUpdate* e passa como parâmetro um objeto do tipo *String* com os dados do sensor. Essa *String* é um JSON (do inglês, Javascript Object Notation).

Como o foco desta aplicação é o movimento do salto, a identificação do salto por acelerômetro é o objetivo do uso do sensor de acelerômetro acoplado ao usuário para descobrir os valores de aceleração nos 3 eixos dimensionais. A partir do padrão de mudança de aceleração durante o movimento do salto, é possível identificar saltos analisando apenas o valor da aceleração do eixo perpendicular à superfície, ou seja, o eixo Y.

5 EXPERIMENTOS REALIZADOS

Para verificar a viabilidade da proposta do VolleyJump, foram realizados testes com um protótipo desenvolvido de acordo com a arquitetura proposta. As tecnologias utilizadas foram: (i) plataforma Android para a aplicação Móvel, (ii) Node.js para o serviço de inferência de saltos, (iii) banco de dados Firebase e (iv) comunicação via Bluetooth LE. Esta seção descreve os experimentos realizados e os resultados obtidos. Como o trabalho também impõe que a coleta dos dados de saltos de um atleta de voleibol de praia a fim de que esta coleta seja feita de maneira imperceptível sem que gere um desconforto ao atleta, foi inicialmente utilizada a *SensorTag CC26505* da *Texas Instruments*, usando uma cinta para acoplá-lo ao peito do atleta, como mostra a figura 4.



Figura 4: Sensortag é acoplado ao peito do atleta a fim de coletar dados durante o salto

Na fase inicial da criação do VolleyJump, para fins de estudo do modelo, foram realizados trinta saltos, entre saltos de ataque e bloqueio, usando a *SensorTag CC26505*. Estes saltos foram coletados no Instituto de Educação Física da Universidade Federal do Ceará (IEFUC), acompanhados por um professor de educação física especializado no treinamento de alto rendimento para atletas de voleibol de praia. Durante a sessão de saltos, os movimentos foram realizados também sobre plataformas de força disponíveis no instituto capazes de informar a altura dos saltos realizados, entre elas a *Jump System Duo*¹, e a *Jump System Optical*², como mostra a figura 5. A amostragem dos valores coletados pelo sensor foi configurada para alta acurácia, a maior possível. Esta decisão segue a orientação de se

¹<http://www.cefise.com.br/produto/104/30/jump-system-duo>

²<http://www.cefise.com.br/produto/78/30/jump-system-optical>

trabalhar com o maior número de amostras possíveis dos sensores de modo a obter maior acurácia da inferência, quando da aplicação dos algoritmos de classificação.



Figura 5: Plataformas utilizadas para validação do trabalho

Durante a análise destes dados, notou-se que os dados gerados pelo dispositivo, ainda que com uma frequência alta de captura de dados, não estavam de acordo com o padrão de dados gerados em um salto como foi mencionado em [7]. A partir de vários experimentos, verificou-se que a *SensorTag CC26505* utilizada em nossos experimentos era muito sensível a variações de seu posicionamento na cinta do atleta. Devido a indisponibilidade de outra *tag*, optou-se por usar o próprio *smart phone Android* acoplado ao peito usando uma cinta como forma melhorar a qualidade dos dados obtidos, uma que de acordo com a arquitetura proposta para o *VolleyJump*, a aquisição dos dados pode ser feita a partir de qualquer dispositivo/sensor capaz de obter dados de um acelerômetro e giroscópio, presos ao corpo do atleta. Os *smart phone Android* utilizados foram o *Motorola Moto E*, *Lenovo Vibe K5* e *OnePlus 5T*. Nesta nova configuração, foram coletados sessenta saltos de um voluntário, divididos entre trinta saltos de ataque e trinta de bloqueio, usando os sensores do celular na altura peito. Estes experimentos tem como objetivo gerar um modelo de classificação do salto, para que ele pudesse ser validado posteriormente. Portanto, cada salto era realizado de maneira discreta, ou seja, o atleta comunicava ao *VolleyJump* o momento em que cada salto era realizado, gerando dados separados por salto.

Com o modelo de classificação gerado, uma nova etapa de coleta de dados foi realizada a fim de validar este modelo e repetir o treinamento. Este experimento foi feito com três estudantes voluntários do curso de Educação Física da Universidade Federal do Ceará, sendo 2 do sexo feminino e 1 do sexo masculino, onde realizou-se uma quantidade total de 127 saltos usando o *VolleyJump* e as plataformas de força mencionadas anteriormente. Estes voluntários representam uma faixa etária de entre 21 e 25 anos, com altura entre 1,80m e 1,85m, pesando entre 80kg e 85kg. Os voluntários não alegaram que o sensor pudesse atrapalhar o desempenho dos seus saltos nem sentiram desconforto ao realizar o salto usando o sensor acoplado ao peito. Os saltos foram realizados usando um protocolo de resistência definido com auxílio do professor de Educação Física que auxilia esta pesquisa, onde o atleta deve saltar o máximo que pode em 5 minutos, com intervalos de 5 segundos entre os saltos. Nesses 5 segundos, o atleta apenas se preparava para realizar o salto seguinte na plataforma, sem que houvesse uma intervenção para

que o *VolleyJump* pudesse identificar este salto, ou seja, os dados coletados neste conjunto de saltos era contínuo, fazendo com que o *VolleyJump* identificasse os saltos. A partir deste experimento, foi possível validar o *VolleyJump* quanto a identificação, cálculo de altura e classificação dos saltos.

5.1 Identificação do Salto

A fim de identificar os saltos realizados, o algoritmo buscava o padrão de aceleração do corpo no decorrer do tempo referente aos saltos mencionado em [7] a cada segundo, pois nos experimentos iniciais notou-se que os saltos teriam uma duração média de 1 segundo, divididos entre tempo no ar e tempo no solo. Com isso, dos 127 saltos realizados no experimento, o algoritmo identificou 129 saltos, gerando um erro de 1,5% de falso positivo.

Percebe-se que o processo de classificação a cálculo de altura do salto depende totalmente da identificação do salto, pois a partir dos saltos identificados que o *VolleyJump* irá classificar e calcular a altura dos saltos realizados, uma vez que, ao identificar o salto, é possível obter o tempo de duração do salto, além dos dados do acelerômetro, gravidade e giroscópio neste intervalo.

5.2 Altura dos Saltos

Em relação à estimativa da altura dos saltos, os valores estimados pelo *VolleyJump* foram comparados com aqueles obtidos pelas plataformas de força. Nas plataformas de força, os saltos tinham 15,27 centímetros de altura em média, enquanto que o *VolleyJump* gerou resultados de 14,06 centímetros, tendo um erro absoluto médio de 1,30 centímetros em relação às plataformas utilizadas. As plataformas de salto são sensíveis ao contato do voluntário com o chão, e isso fez com que 14 dos 127 saltos fossem ignorados dos cálculos, pois a plataforma não mostrava dados reais.

5.3 Classificação dos Saltos

Após a coleta dos 60 saltos realizados no primeiro experimento válido, o algoritmo gerou um modelo de classificação destes saltos usando um algoritmo de máquinas de vetores de suporte, a fim de que, no momento da coleta dos dados, não fosse mais necessário que o jogador classifique os saltos sempre que a ação for executada. Para a geração do modelo, os dados foram ordenados de maneira aleatória e separados entre dados de treino e validação do modelo, onde 70% dos dados foram usados para treinar o modelo e os 30% restantes foram usados para a validação. Para evitar que o modelo ficasse tendencioso, os saltos separados para a realização do treino do modelo estavam divididos igualmente entre os saltos de ataque e bloqueio.

A figura 6 ilustra o padrão dos dados do acelerômetro e giroscópio no eixo Y na realização de uma sequência de 5 saltos. É possível perceber o padrão da identificação do salto caracterizado pela queda brusca de aceleração para tomar impulso e recuperação da aceleração ao finalizar o salto. No giroscópio, é importante perceber o intervalo de valores resultantes durante o salto, variando entre -0.9 e 0.9.

A fim de comparar os saltos, a figura 7 ilustra os dados coletados durante 6 saltos. O padrão de identificação do salto permanece ao analisar os dados do acelerômetro e o giroscópio tende a ser mais variável, variando entre aproximadamente 2 e -2. Essa maior

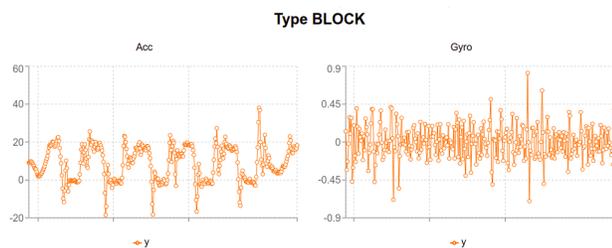


Figura 6: Acelerômetro e giroscópio no eixo Y no momento de uma sequência de saltos de bloqueios

variância do giroscópio no eixo Y ocorre porque o salto de ataque é caracterizado por uma por um avanço prévio ao salto, realizando uma passada antes do salto, diferente do salto de boqueio, onde o atleta salta praticamente partindo do mesmo lugar em que se encontra. Os dados coletados nos eixos X e Z não representaram um padrão que pudesse ser relevante às diferenças de ataque entre salto e bloqueio.

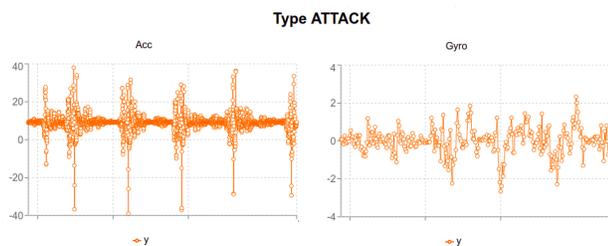


Figura 7: Acelerômetro e giroscópio no eixo Y no momento de uma sequência de saltos de ataque

No último experimento realizado, um conjunto de 85 saltos de bloqueio e 42 saltos de ataque foram executados. A tabela 1 resume os resultados da classificação dos saltos a partir destes saltos.

Tabela 1: Dados coletados durante experimentos com voluntários

	Classif. como Bloqueio	Classif. como Ataque
Bloqueio	78	5
Ataque	12	32
Total	90	37

Na validação do modelo, obteve-se uma taxa de acerto de 86,61% a partir do experimento realizado, onde 93,97% dos saltos de bloqueio e 72,72% dos saltos de ataque foram corretamente classificados. O resultado apresentado é promissor, porém observa-se a necessidade de ajustar o modelo pra que reduza o numero de saltos de ataques classificados erroneamente como saltos de bloqueio. Uma possível explicação para este fenômeno está na menor quantidade de saltos de ataque utilizados para construir o modelo de classificação. O desbalanceamento entre as classes (saltos de ataque e saltos de bloqueio) tende a construir modelos que privilegiam a classe majoritária (com mais exemplos). Desta forma, a continuação deste

trabalho prevê uma maior coleta de dados de ataque para reduzir o impacto do desbalanceamento no desempenho do modelo.

6 CONCLUSÃO

O uso de tecnologias digitais é uma tendência cada vez maior em esporte de alto rendimento. Este trabalho apresentou como técnicas de aprendizagem de máquina podem ser utilizadas para classificar saltos de voleibol de praia, e também inferir a altura destes saltos a partir de dados de acelerômetro gravidade e giroscópio presos ao corpo de um atleta. Este trabalho apresenta um aplicativo chamado *VolleyJump*, desenvolvido para a plataforma Android, que apresenta-se como uma alternativa mais econômica a análise do desempenho dos atletas de voleibol de praia.

O *VolleyJump* mostrou que é viável realizar análises de movimentos de atleta sem um alto custo financeiro. A aplicação pretende ser distribuída na forma de software livre sobre a licença Apache 2.0, a fim de que estudantes, profissionais da Educação Física e equipes de voleibol de praia amadoras e profissionais possam ter acesso a uma ferramenta de baixo custo com resultados que possam viabilizar a melhoria de rendimento entre atletas. Como trabalho futuro a fim de aumentar o conforto do atleta, deseja-se criar um protótipo de um sensor vestível ao atleta sem que seja necessário acoplar um dispositivo móvel ao seu corpo. O código-fonte do *VolleyJump* pode ser visto em <https://github.com/renanbandeira/VolleyJump>.

REFERÊNCIAS

- [1] Ralph Johnson Richard Helm Erich Gamma, John Vlissides. 1994. *Design Patterns: Elements of Reusable Object-Oriented Software*.
- [2] Mohammed Abdulaziz Ikram, Mohammad Dahman Alshehri, and Farookh Khadeer Hussain. 2015. Architecture of an IoT-based system for football supervision (IoT Football). *2015 IEEE 2nd World Forum on Internet of Things (WF-IoT)* (2015), 69–74.
- [3] Jon M Jarning, Kam-Ming Mok, Bjoerge Herman Hansen, and Roald Bahr. 2015. Application of a tri-axial accelerometer to estimate jump frequency in volleyball. *Sports biomechanics* 14 1 (2015), 95–105.
- [4] Thomas Kautz, Benjamin H. Groh, Julius Hannink, Ulf Jensen, Holger Strubberg, and Bjoern M. Eskofier. 2017. Activity recognition in beach volleyball using a Deep Convolutional Neural Network. *Data Mining and Knowledge Discovery* 31, 6 (01 Nov 2017), 1678–1705.
- [5] Nicholas P. Linthorne. 2001. Analysis of standing vertical jumps using a force platform. *American Journal of Physics* 69 (2001).
- [6] Tiago Ribeiro, Octavian Postolache, and Pedro Passos. 2016. Performance assessment for mountain bike based on WSN and cloud technologies. *2016 International Conference and Exposition on Electrical and Power Engineering (EPE)* (2016), 380–386.
- [7] Lilian K. S. Sunna. 2010. *Athlete monitoring via wireless sensor on a mobile phone*. Master's thesis.
- [8] Jan Van Haaren, Horesh Ben Shitrit, Jesse Davis, and Pascal Fua. 2016. Analyzing Volleyball Match Data from the 2014 World Championships Using Machine Learning Techniques. *Conference: the 22nd ACM SIGKDD International Conference* (08 2016), 627–634.

QuiRA: An Augmented Reality Application to Support Chemistry Learning

Letícia Porto Soares
Instituto Federal da Bahia
Vitória da Conquista, Brasil
leticiaportosoares@gmail.com

Camilo Alves Carvalho
Instituto Federal da Bahia
Vitória da Conquista, Brasil
camilocarvalho@ifba.edu.br

Luis Paulo da Silva Carvalho
Instituto Federal da Bahia
Vitória da Conquista, Brasil
luispscavalho@gmail.com

Manuela Amaral de Araújo
Instituto Federal da Bahia
Vitória da Conquista, Brasil
mn21922192@gmail.com

Paulo Henrique Teixeira Silva
Instituto Federal da Bahia
Vitória da Conquista, Brasil
paulohenrique-201119@hotmail.com

Wdson Costa Santos
Instituto Federal da Bahia
Vitória da Conquista, Brasil
wdsoncosta@ifba.edu.br

ABSTRACT

The volume of information which contemporary student has access is huge, and they are surrounded by smartphones, tablets, internet and computers in an almost inseparable way from their daily life. Therefore, we think that it is important that schools should consider the insertion of technological instruments in the classroom to attend the students' interest in these devices. One of the alternatives is combining Augmented Reality and mobile technology to enable the interaction with virtual objects by overlapping them with real environment. These technologies may become allied in the process of teaching and learning abstract concepts of various subjects. In this context, QuiRA is proposed to help the teaching of Chemistry, an ideal field to explore questions about the use of computer animation due to the difficulty of visualizing its concepts in two dimensions. This article discusses the details about the QuiRA application and the demonstration of its use for visualizing chemical molecules.

KEYWORDS

Augmented Reality, ARToolkit. Chemistry.

1 INTRODUÇÃO

A natureza do conhecimento está em constante evolução. No entanto, os processos de ensino e aprendizagem contemporâneos são considerados obsoletos. Hoje, o volume de informações é imenso e o aluno se vê envolto por tecnologias como: *smartphones*, *tablets*, *internet*, computadores, entre outros, e como estão muito presentes, quase que indissociáveis ao cotidiano deles, fornecem experiências mais atrativas. De acordo com [1], cabe, então, aos educadores utilizarem instrumentos tecnológicos em sala de aula, para atender não apenas um interesse exclusivamente pedagógico, mas os interesses dos alunos.

Uma das alternativas que utiliza da tecnologia para uma aprendizagem mais lúdica e efetiva é a inserção da Realidade Aumentada (RA) que, segundo [2], se caracteriza por combinar real com virtual, possuir interatividade em tempo real e ser registrado em três dimensões. A RA se trata de uma tecnologia onde os alunos poderão interagir com objetos virtuais através da sobreposição de camadas virtuais a um ambiente real [3]. Essa interação de ambientes virtuais no mundo real torna a percepção mais atraente e estimula o processo investigativo dos alunos. Considerando os avanços da tecnologia móvel e a popularização dos smartphones é possível integrar a Realidade Aumentada (RA) aos dispositivos móveis inserindo um novo paradigma à relação aluno-professor, onde a utilização desses dispositivos se torna aliada, e não um incômodo, no processo de ensino e aprendizagem de diversas disciplinas, como, química, matemática, biologia, geografia e história.

O QuiRA foi proposto para auxiliar o ensino de Química por meio da Realidade Aumentada aliada aos dispositivos móveis. O objetivo da aplicação foi direcionado para a Química, porque: (i) é um componente curricular obrigatório no ensino básico, (ii) é considerada de difícil compreensão pelos alunos e, (iii) por muitas vezes, os alunos não conseguem associar o conteúdo da disciplina ao seu dia-a-dia [4] e (iv) pela visualização em 2D não ser suficiente para a devida observação do arranjo tridimensional dos átomos nos compostos químicos [5], sendo, portanto, um conteúdo propício para explorar as potencialidades da Realidade Aumentada.

Portanto, esse trabalho tem o objetivo de trazer uma ferramenta de Realidade Aumentada, que no âmbito do ensino, proporcione aos educadores uma maneira agradável de transmitir os conteúdos aos alunos, a fim de estimular a aprendizagem. Além disso, no âmbito técnico, pretende-se apresentar uma abordagem de utilização da biblioteca de Realidade Aumentada ARToolkit em dispositivos móveis Android que dê ao usuário a possibilidade de escolher quais objetos tridimensionais ele deseja visualizar através da RA e armazená-los em um banco de dados. Este fato potencializa e colabora com a possibilidade de criação de um visualizador de RA genérico com uma biblioteca de acesso gratuito.

2 TRABALHOS RELACIONADOS

Em [6], Cunha e Guimarães criaram uma ferramenta gráfica que tem o objetivo de tornar o desenvolvimento de aplicações de RA mais fácil e transparente ao usuário, neste caso professores e alunos. No seu desenvolvimento foi utilizado o software Microsoft Visual Basic 6.0 com a biblioteca *ARToolKit*¹. A ferramenta permite aos usuários a associação dos objetos virtuais aos marcadores predefinidos, não sendo permitida a inserção de novos marcadores. O seu funcionamento se dá através da sua execução em um computador de mesa com uma *webcam* para detectar o marcador e permitir a visualização do objeto virtual projetado.

O Serviço Nacional de Aprendizagem Industrial (SENAI) produziu aplicativos móveis que utilizam realidade aumentada no programa SENAI de Tecnologias Educacionais [7]. Com os aplicativos, os alunos podem utilizar a câmera dos dispositivos móveis, do celular ou do *tablet*, para reconhecimento em 3D especificamente das imagens presentes nos livros didáticos do SENAI. Além disso, os aplicativos permitem aos estudantes acessarem simuladores e vídeos, pelos quais os objetos impressos ganham movimento, imagem, som e interação.

Em [8] foi efetuado um experimento com o desenho e construção de objetos de aprendizagem baseados em realidade aumentada, voltado para alunos do ensino secundário de uma escola na Colômbia. Na ferramenta de RA foram utilizados os softwares *Unity*, *Vuforia* e *Android Studio* para a implementação e o *3ds Max* para construção de 35 modelos em 3D. Foram produzidas imagens em duas dimensões para serem referência para os 35 modelos tridimensionais construídos. Após o estudo, foram obtidas avaliações positivas em termos: (a) de satisfação de uso; (b) com as características estéticas, funcionais e de conteúdo dos objetos; (c) e com realidade aumentada dentro das salas de aula como motivação para aprender. Os professores da escola, inclusive, destacaram a capacidade de foco dos alunos nas atividades de interação com o software.

Em [9] foi produzida a ferramenta *Reações Químicas Aumentadas (Augmented Chemical Reactions)* que visa auxiliar tanto alunos quanto cientistas mostrando moléculas químicas renderizadas em 3D quando um marcador é direcionado a câmera do computador. Além disso retrata a alteração dinâmica de moléculas quando se aproximam. Neste caso, usuários podem ter uma visão melhorada sobre certos comportamentos e interações entre moléculas. Para o seu desenvolvimento foram utilizados *ARToolKit* e *UbiTrack*. A ferramenta aborda duas formas de utilização. A primeira voltada aos alunos, onde a configuração é feita previamente e estes só irão utilizá-la para visualização das moléculas. E a segunda, onde, num modo especialista, o usuário irá definir as moléculas e atribuí-las a marcadores dinamicamente.

O QuiRA traz uma abordagem dinâmica no que se diz respeito à possibilidade de inclusão, pelo usuário, de novos objetos tridimensionais a aplicação, o que o diferencia dos trabalhos já existentes. As ferramentas disponíveis mais atuais

proporcionam um conteúdo estático de marcadores e/ou objetos tridimensionais. Já as aplicações que permitem que sejam adicionados novos objetos 3D e associá-los a marcadores foram desenvolvidas para computadores de mesa munidos de uma *webcam*. Devido a essas características elas podem ser consideradas inadequadas diante do advento da computação móvel, da popularização dos dispositivos móveis e da maior viabilidade proporcionada pela utilização de *smartphones* e *tablets* em comparação com dispositivos antigos, como, por exemplo, os computadores de mesa.

3 QUIRA

O QuiRA é um aplicativo móvel baseado em Realidade Aumentada, desenvolvido com o objetivo de auxiliar no processo de ensino e aprendizagem da disciplina Química, em especial do conteúdo de Química Orgânica no contexto da Isomeria. O software permite que o usuário cadastre moléculas químicas para visualização em três dimensões (3D). A projeção dos objetos em 3D tem como referência os marcadores monocromáticos. Esses marcadores são constituídos por uma margem branca, borda preta e um identificador em seu interior, um símbolo ou uma letra, de fácil identificação no contexto computacional [10], conforme o exemplo mostrado na Figura 1.

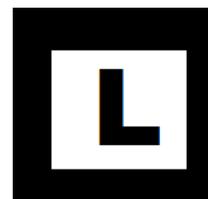


Figura 1. Marcador monocromático²

3.1 Arquitetura do QuiRA

A arquitetura do QuiRA (conforme exibida na Figura 2) possui os componentes *ActCadObjeto*, *ActEdtObjeto* e *ActListaObjetos* para as operações de inclusão, edição, listagem e exclusão das moléculas em uma tabela denominada Objeto do banco de dados. Quando o dispositivo de câmera é selecionado dentro do aplicativo, a biblioteca *ARToolKit* captura a imagem obtida pela câmera e envia para o componente *ARWrapperObjeto*. Este componente: (i) identifica o marcador presente na imagem recebida e (ii) tem acesso aos objetos cadastrados no banco de dados. Quando *ARWrapperObjeto* reconhece que o marcador rastreado está associado a um objeto cadastrado no banco de dados, desenha na tela do dispositivo móvel o objeto em três dimensões sobre ao marcador.

A seguir se encontram detalhadas as tecnologias e recursos mais importantes utilizados pelos componentes da arquitetura do QuiRA.

¹ <https://github.com/artoolkit>

² Este marcador poderá ser utilizado posteriormente para testar o aplicativo

3.2 ARToolkit

A realidade aumentada foi implementada na aplicação por meio da biblioteca *Augmented Reality Toolkit (ARToolkit)*, uma biblioteca de visão computacional que fornece a funcionalidade de rastreamento necessária para criar aplicativos de realidade aumentada [11]. O uso do *ARToolkit* requer o relacionamento entre componentes escritos em linguagem de programação Java e funções em C++. Para tanto, fez-se uso da estrutura *Java Native Interface (JNI)* que permite essa comunicação [12] e do *Android Native Development Kit (NDK)*, ferramenta do *Android* que fornece componentes para utilizar o JNI [13]. O JNI foi empregado no contexto do QuiRA, porque este recurso melhora o desempenho da ferramenta.

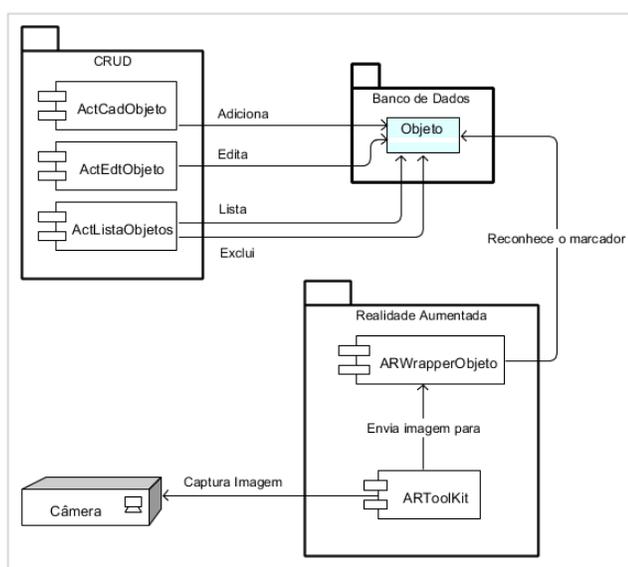


Figura 2. Arquitetura do QuiRA

O *ARToolkit* utiliza dois arquivos para a projeção dos objetos. O primeiro é a digitalização do marcador monocromático exportada para um arquivo no formato .patt. Para essa digitalização, a própria biblioteca fornece a aplicação *Mk_Patt*. O segundo arquivo é o objeto em três dimensões que deve ser exportado para o formato .obj. O arquivo .obj (OBJ) pode ser gerado através de softwares de desenho gráfico, como o *Blender*³, por exemplo.

3.3 Compostos

Para a utilização no QuiRA foram produzidos treze compostos químicos através do software *Blender* exportando-os para o formato .obj. Os compostos foram ácido etanoico, metanoato de metila, etanol, metoximetano, propanal,

propanona, ácido propanoico, etanoato de metila, butan-1-ol, etoxietano, but-2-en-1-ol, butan-2-ona e butanal. Na figura 3, observa-se o composto Propanal gerado.

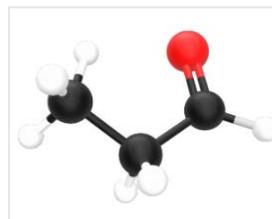


Figura 3. Propanal

3.4 Implementação

O processo de desenvolvimento do *software* envolveu: (a) a definição do banco de dados para armazenamento das moléculas e sua implementação; (b) o desenvolvimento da interface para cadastro; (c) e gerenciamento das moléculas e expansões nas classes de realidade aumentada.

Para o armazenamento das moléculas cadastradas no aplicativo, foi projetado um banco de dados no *SQLite*, biblioteca que implementa um banco de dados SQL e que é nativa do *Android*. No *SQLite* foi definida a tabela denominada *Objeto*, cujos atributos são mostrados na Tabela 1.

Para o desenvolvimento da interface do aplicativo, definiu-se uma tela inicial com um botão de câmera para acessar a tela de visualização dos objetos 3D e com um menu com as opções de gerenciamento de objetos, de acesso ao *Sobre* do aplicativo e ao *Manual* do aplicativo. A tela de Cadastro de objeto permite a inserção do nome, descrição, arquivo .patt, arquivo .obj, imagem do objeto e imagem do marcador. Uma tela de listagem foi disponibilizada para selecionar objetos. Essa seleção permite visualizar o cadastro de um objeto. Além da visualização, um objeto pode ser modificado e excluído.

Table 1: Atributos da tabela Objeto

Atributo	Armazena
CÓDIGO	Código identificador
NOME	Nome da molécula
DESCRICAÇÃO	Observação sobre a molécula
PATT	Caminho do arquivo .patt gerado através do marcador que será referência para a projeção da molécula
OBJ	Caminho do arquivo .obj da molécula que será renderizado na tela
IMAGEM	Caminho da imagem da molécula
MARCADOR	Caminho da imagem do marcador

Cada tela que pode ser visualizada pelo usuário foi definida em código com uma *Activity*, como mostrado na Figura 4. É possível notar que a partir da tela Início é possível atingir as demais e que a tela de listagem se torna o meio através da qual

³ <https://www.blender.org/>

são utilizadas as telas de inclusão, edição e exclusão de objetos correspondentes a moléculas.

A Realidade Aumentada também foi inserida no aplicativo amparada por uma *Activity*. Essa *Activity* efetua comunicação com o banco de dados para recuperar os caminhos dos arquivos .patt e .obj gravados através da tela de Cadastro.

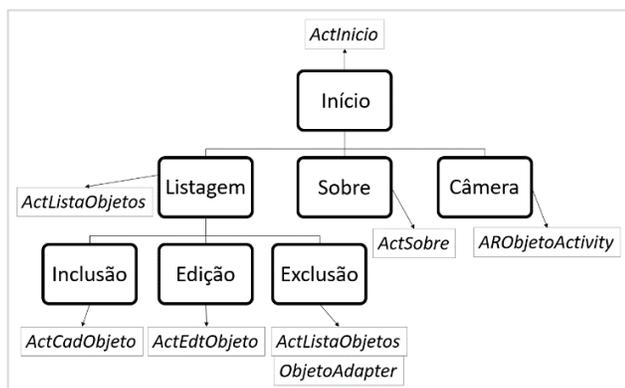


Figura 4. Telas e suas respectivas *Activities*

4 UTILIZAÇÃO DO QUIRA

O aplicativo QuiRA pode ser utilizado através da sua instalação em um dispositivo móvel com sistema operacional *Android*. Os materiais necessários para a sua instalação e utilização estão disponíveis no endereço bit.ly/2w8EQWA. O endereço bit.ly/2nQP8qF contém um vídeo que demonstra esse processo.

Através do endereço bit.ly/2w8EQWA citado, pode-se efetuar: (i) o *download* do arquivo instalador do aplicativo, (ii) obter os arquivos para cadastro da molécula química Butanal, (iii) acessar o README para obter passos rápidos para a utilização do QuiRA, (iv) visualizar o manual que ampara o processo de utilização do aplicativo de uma forma mais completa, desde a criação de marcadores até a visualização de objeto de 3D, e (v) adquirir o Digitalizador de Marcadores Mk_Patt.

Após a instalação do QuiRA, a sua tela inicial (Figura 5a) e o seu menu podem ser acessados. Com o cadastro da molécula Butanal (Figura 5b), esta pode ser visualizada através do módulo de câmera da aplicação quando a câmera do dispositivo é direcionada ao marcador relacionado. Para este caso foi cadastrado o marcador com a letra L. A medida que novas moléculas são cadastradas na aplicação, elas são listadas como na Figura 6a. O cadastro de cada molécula pode ser acessado individualmente, o que possibilita sua visualização completa e sua edição (Figura 6b). A molécula pode ser excluída da listagem por meio do botão de lixeira. Todas as moléculas cadastradas serão carregadas no momento que o módulo de câmera for acionado. Com isso elas poderão ser visualizadas em três dimensões conforme mostrado na Figura 7. E a visualização do objeto em 3D atende ao movimento do marcador.

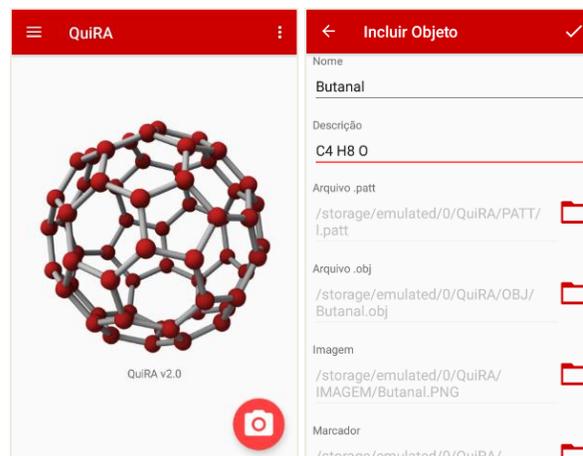


Figura 5. Telas Inicial e de Cadastro

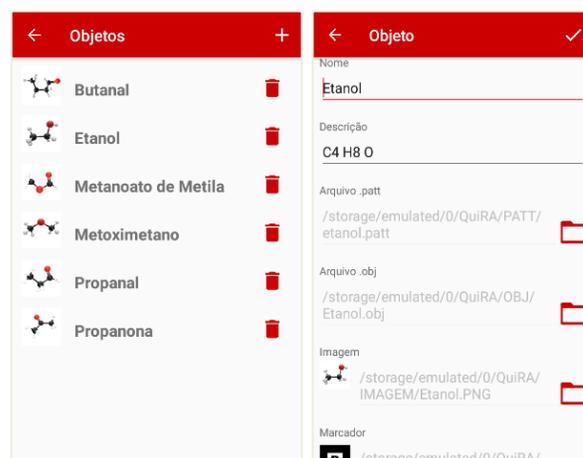


Figura 6. Telas de Listagem e de Edição

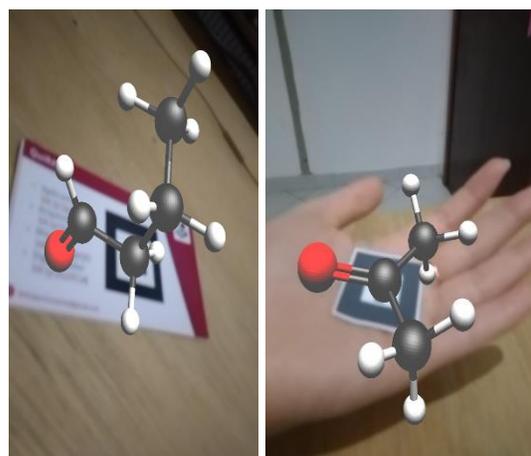


Figura 7. Visualização das moléculas Butanal e Propanona

4 CONCLUSÕES

A Realidade Aumentada pode trazer uma experiência diferenciada ao usuário e o QuiRA permite que essa tecnologia seja trazida para mais perto do usuário em um alto nível de abstração. No aplicativo é possível ter a vivência com a realidade aumentada sem a necessidade de lidar com programação de códigos, tornando o contato com a RA mais acessível. Dessa forma, o professor pode utilizá-lo em sala de aula como incentivo para os alunos, já que possibilita que conteúdos intangíveis e de difícil aplicação prática, possam ser visualizados através de dispositivos que já fazem parte de seu cotidiano. O aplicativo permite que o usuário efetue o cadastro das moléculas, associando sua representação tridimensional a um marcador monocromático, a partir de arquivos armazenados no dispositivo móvel. Depois, as moléculas podem ser vistas instantaneamente através da câmera do dispositivo dentro do aplicativo, quando esta é direcionada aos marcadores cadastrados. Um trabalho futuro inclui o desenvolvimento, a partir do QuiRA, de aplicativos que amparem outras disciplinas, a fim de colaborar com uma aprendizagem mais lúdica. Outra possibilidade é desenvolver aplicativos direcionados a visualização de objetos de outras áreas, como a Arquitetura. Diante disso, o QuiRA é um projeto com pretensão de evolução e que ainda está em desenvolvimento.

Além de abordar outras áreas de ensino, também existe a intenção de avaliar o QuiRA em um experimento. Tal experimento será realizado com o objetivo de coletar opiniões de professores e alunos. Posteriormente, essas opiniões serão utilizadas para melhorar o QuiRA tanto em relação à sua usabilidade quanto, possivelmente, à sua arquitetura e codificação. Por isso, por enquanto, a sua licença permanece fechada, mas a intenção é que o seu código fonte seja liberado.

REFERÊNCIAS

- [1] O. H. A. Araújo, L. T. F. Ribeiro e M. N. S. Pinheiro. Tecnologias móveis nos processos de ensino e de aprendizagem: mobilidade docente?. Revista Ibero-Americana de Estudos em Educação, v. 11, n. 1, p. 95-110, 2016, in press. Disponível em: <<https://periodicos.fclar.unesp.br/iberoamericana/article/view/8087>> Acesso em 10 Maio 2018
- [2] R. Azuma. Recent advances in augmented reality. In v. 21, n., editor, IEEE Computer Graphics and Applications, pages 34-47, in press. 2001. Disponível em: <<http://www.cs.unc.edu/~azuma/cga2001.pdf>> Acesso em 29 Maio 2018
- [3] L. Araujo, E. Xavier, P. Vasconcelos, M. Machado e T. Tavares (2017). DoctorBio: Um estudo de caso sobre a utilização de recursos de realidade aumentada no ensino de ciências biológicas. In: Anais do Workshop de Informática na Escola, in press. 2017. p. 294. Disponível em: <<http://www.br-ie.org/pub/index.php/wie/article/view/7247/5045>> Acesso em 10 Maio 2018
- [4] A. L. Marques et al (2008). A importância de aulas práticas no ensino de química para melhor compreensão e abstração de conceitos químicos. In: Encontro Nacional de Ensino de Química, 14., 2008, Paraná. Anais... Paraná: UFPR, 2008, in press. Disponível em <<http://www.quimica.ufpr.br/eduquim/eneq2008/resumos/R0727-1.pdf>> Acesso em 10 Maio 2018.
- [5] ARAÚJO, M. A (2017). A Realidade Aumentada no contexto educacional como mecanismo alternativo para visualização de Isômeros na Química Orgânica. Monografia (Graduação em Sistemas de Informação). Instituto Federal da Bahia, Vitória da Conquista, 2017.
- [6] K. K. C. Cunha, M. P. Guimarães. Desenvolvimento de uma Interface Gráfica para o Artoolkit com Aplicação na Área Educacional. In: WRVA'07 Workshop de Realidade Virtual e Aumentada, Itumbiara, 2007, p. 1-4, in press. Disponível em <<http://www.lbd.dcc.ufmg.br/colecoes/wrva/2007/0030.pdf>> Acesso em 10 Fev 2018.
- [7] PORTAL DA INDUSTRIA. Alunos do SENAI usam aplicativos de realidade aumentada para aprender de forma divertida e inovadora. 2016. Disponível em <<http://www.portaldaindustria.com.br/agenciadni/noticias/2016/08/alunos-do-senai-usam-aplicativos-de-realidade-aumentada-para-aprender-de-forma-divertida-e-inovadora/>> Acesso em 10 Fev 2018.
- [8] E. J. Hernandez-Leak. Construction of learning objects with Augmented Reality: An experience in secondary education. In: Learning Technologies (LACLO), 2017 Twelfth Latin American Conference on. IEEE, 2017. p. 1-7, in press. Disponível em <<https://ieeexplore.ieee.org/document/8120948/authors>> Acesso em 12 Maio 2018.
- [9] P. Maier, M. Tönnis, G. Klinker. Augmented Reality for teaching spatial relations. In: Conference of the International Journal of Arts & Sciences, Toronto, 2009, in press. Disponível em <<https://pdfs.semanticscholar.org/6813/b691bc09e0172051b2e77c379d74b930149c.pdf>> Acesso em 10 Fev 2018.
- [10] ANDROID PRO. Android SDK. 2017. Disponível em <<https://www.androidpro.com.br/blog/android-studio/android-sdk/>> Acesso em 29 Maio 2018.
- [11] GITHUB. android_about.md. 2016. Disponível em <https://github.com/artoolkit/artoolkit-docs/blob/master/4_Android/android_about.md> Acesso em 12 Maio 2018.
- [12] GITHUB. android_developing.md. 2016. Disponível em <https://github.com/artoolkit/artoolkit-docs/blob/master/4_Android/android_developing.md> Acesso em 12 Maio 2018.
- [13] ANDROID. Primeiros passos com o NDK. Disponível em <<https://developer.android.com/ndk/guides/index.html?hl=pt-br>> Acesso em 15 Ago 2017.

MCU em software distribuído em nuvem

Valter Roesler
Instituto de Informática, UFRGS –
Universidade Federal do RGS
Porto Alegre, RS, Brasil
roesler@inf.ufrgs.br

Mário Gasparoni Júnior
Instituto de Informática, UFRGS –
Universidade Federal do RGS
Porto Alegre, RS, Brasil
mariogasparoni@gmail.com

Ronaldo Husemann
Instituto de Informática, UFRGS –
Universidade Federal do RGS
Porto Alegre, RS, Brasil
rhusemann@inf.ufrgs.br

Roberto Irajá T. da Costa Filho
Instituto de Informática, UFRGS –
Universidade Federal do RGS
Porto Alegre, RS, Brasil
rtcosta@gmail.com

Rafael Valle
RNP – Rede Nacional de Ensino e
Pesquisa
Rio de Janeiro, RJ, Brasil
rafael.valle@rnp.br

ABSTRACT

The presented tool is a fully virtualized videoconferencing MCU (*Multipoint Control Unit*) system using the standard SIP (*Session Initiation Protocol*). The proposed tool works in the cloud in a scalable way, with low deployment and maintenance costs. In addition, the proposed tool is more than an MCU, functioning as a universal framework for media forwarding.

KEYWORDS

Multimedia; MCU; SFU; videoconference; web conference; video collaboration.

1 INTRODUÇÃO

Hoje em dia as instituições utilizam soluções de videoconferência e videoconferência como uma forma de redução de custos operacionais (viagens, por exemplo) e também aumento na velocidade na qual a interação é efetuada. Atualmente o mercado vê um crescimento no uso de videoconferência, provocado muito pelo declínio do custo de ferramentas de videoconferência baseadas em software, como Microsoft Skype for Business e WebEx, que aproveitam equipamentos já existentes na empresa, como laptops, tablets e smartphones (equipados com câmeras). Essas soluções em software custam significativamente menos que as equivalentes em hardware, ampliando o acesso [1]. O modelo de salas de videoconferência onde várias pessoas participam juntas da interação é muito utilizado atualmente, em paralelo com o modelo de uso em dispositivos pessoais. O motivo é que, muitas vezes, a interação em um ambiente único em cada ponto facilita a comunicação. O fato é que existem diversos modelos atualmente para efetuar videoconferência, e diversas ferramentas para isso [2].

Ferramentas de videoconferência são úteis em comunicações ponto a ponto, com dois interlocutores, e também em

comunicações multiponto, com vários interlocutores. Para comunicações multiponto, um dos modelos existentes é o de MCU (*Multipoint Control Unit*), que permite a comunicação entre diversos participantes numa videoconferência, com fluxo de saída composto pelas mídias de diversos equipamentos finais de videoconferência (também chamados de *endpoints*) [3].

Os *endpoints* podem ser compostos por sistemas de hardware ou de software. Exemplos de *endpoints* de hardware são a linha HDX da Polycom [4], o Multipoint Bridge da Lifesize [5], e a linha MX da Cisco [6], entre outros. Exemplos de *endpoints* de software são o Jitsi [7], o Ekiga [8] e o RealPresence da Polycom [9].

Para interconectar os diversos *endpoints*, um MCU age como elemento centralizador, recebendo todas as mídias, compondo todas em um mesmo vídeo, e enviando esse vídeo para todos *endpoints*, como mostra a Figura 1. Observa-se que, em termos de rede, o MCU recebe um fluxo de mídia (áudio + vídeo) de cada usuário, e envia um fluxo de mídia composta (áudio mixado + vídeo composto) para cada participante.

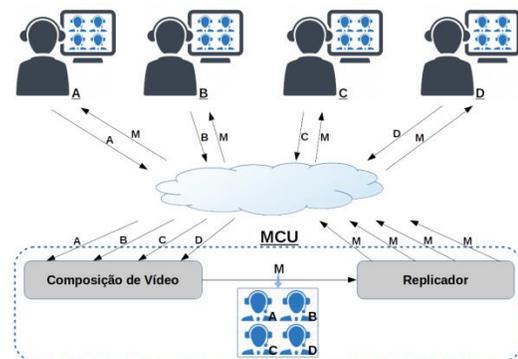


Figura 1: Funcionamento básico de um MCU

Essa arquitetura centralizada onera o servidor MCU, que precisa processar (decodificar e/ou transcodificar) os fluxos de todos *endpoints* e compor (codificar) o fluxo de saída. Devido à elevada carga de processamentos de vídeo que um MCU necessita atender, os equipamentos de MCU são produtos

bastante especializados, compostos normalmente por hardwares de DSPs (*Digital Signal Processors*) ou por módulos específicos de software instalados em servidores COST (*commercial off-the-shelf*), especialmente disponíveis e dedicados à decodificação e codificação de vídeo e áudio. Essa complexidade torna os produtos comerciais de MCU bastante caros para aquisição, e também caros para manutenção [10].

É possível fazer MCUs mais baratos, em software, e existem várias iniciativas para isso, como o projeto OpenMCU-ru (<https://videowitch.ru/eng.htm>), que apresenta uma plataforma completa em código aberto de um MCU. O problema de MCUs em software é que, por rodar em cima de um processador não especializado (genérico), o número de *endpoints* que uma máquina consegue atender é limitado, menor do que as soluções dedicadas, conforme será detalhado adiante neste documento. Esse fato se agrava quando se trabalha com qualidade HD.

A proposta deste artigo é apresentar um sistema de MCU em software que permite escalabilidade na nuvem, mantendo o custo de aquisição e manutenção baixo, por se tratar de software, mas permitindo escalabilidade, suportando um número muito grande de usuários, de acordo com o número de máquinas virtuais alocadas para o sistema.

A seção 2 apresenta uma visão geral da ferramenta, mostrando seus diferenciais. A seção 3 mostra sua arquitetura interna, detalhando ainda mais seu funcionamento. A seção 4 mostra alguns trabalhos relacionados. As seções 5 e 6 mostram o ambiente experimental de validação e os resultados obtidos.

2 ARQUITETURA GERAL DA FERRAMENTA

A Figura 2 mostra a arquitetura geral da ferramenta proposta, que é composta por um Gerenciador de Escalabilidade (GE) redundante e uma rede de Servidores de Mídia (SM) de MCUs em software, distribuídos na nuvem. A redundância do GE é necessária para que o sistema possua alta disponibilidade, visto que o GE gerencia toda a sinalização, e possui a inteligência de todas as conferências em andamento, tornando-se um elemento crítico no sistema. A redundância dos Servidores de Mídia é necessária para, além de fornecer alta disponibilidade, aumentar o poder de processamento do sistema, suportando mais *endpoints*. Os *endpoints* conectam com o GE (linhas pretas na figura), que é, basicamente, o ponto de entrada do MCU para os usuários. O GE escolhe um Servidor de Mídia disponível e direciona este usuário (*endpoint*) para ingresso na sala desejada, fazendo com que toda a transferência de mídia (vídeo e áudio) seja feita agora diretamente entre SM e *endpoint* (ver figura). A figura também representa várias salas independentes com vários usuários em cada sala, todos se comunicando através do MCU proposto.

Em termos de características gerais, a ferramenta contempla:

1. Funcionamento em software, em máquina virtual, de forma distribuída em nuvem.
2. Escalável através de um gerenciador de escalabilidade, permitindo diversas conferências simultâneas.
3. Baixo custo de implantação e de manutenção.

4. Funcionamento como um Servidor de Mídia Universal na rede, permitindo receber demandas de envio de áudio e vídeo e encaminhar para o(s) destino(s) desejado(s) no formato definido.
5. Alta robustez e disponibilidade para uso contínuo, ou seja, a ferramenta passou por baterias de testes de carga e funcionamento contínuo, provendo para o usuário uma experiência funcional, sem interrupções ou reinicializações.

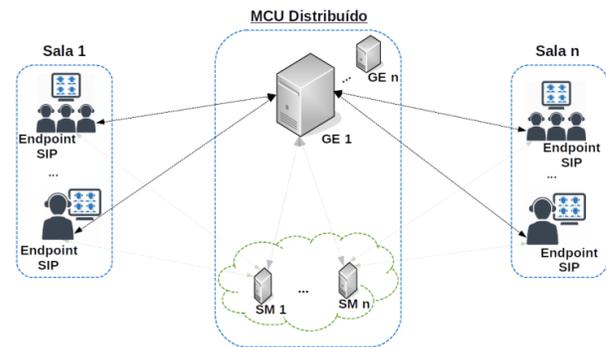


Figura 2: Visão geral do sistema MCU em nuvem.

3 ARQUITETURA INTERNA DO MCU

A Figura 3 exibe a arquitetura do MCU, mostrando os principais blocos do Gerenciador de Escalabilidade, resumidos a seguir:

- **MCS-SIP (*Media Control Server – Session Initiation Protocol*)**: atua como receptor das requisições SIP, traduzindo-as em chamadas de API para o servidor MCS. Além de traduzir as mensagens SIP para a API própria, esse bloco envia mensagens *SIP Redirect* para o *endpoint*, a fim dele trocar a mídia diretamente com o SM designado.
- **MCS (*Media Control Server*)**: corresponde a uma aplicação e um framework para desenvolvimento de aplicações multimídia. Uma de suas principais funções é fazer o balanceamento de carga dos servidores de mídia disponíveis. Além disso, o MCS define um modelo de abstração que é independente do tipo de comunicação de mídia realizado (RTP, WebRTC, etc.) permitindo a gerência de salas, usuários e mídias. Dessa forma, um usuário de uma webconferência (acessando uma sala de MCU através do navegador WEB de um PC) é capaz de se comunicar com um usuário de videoconferência (acessando a mesma sala de MCU através de um *endpoint* de videoconferência).
- **Gerente**: através de uma interface web, o Gerente permite o acesso Federado para controle do sistema. Esse módulo permite a criação de salas, desde que o usuário tenha permissão para isso. Além disso, cada sala tem um gerente local (moderador), responsável por mudanças de layout, início de gravação, expulsar um usuário, colocar em mudo, entre outras. Esse módulo também permite a gerência local da sala. A comunicação com o MCS é através da mesma API mencionada anteriormente.
- **BD Distribuído**: todos os dados instantâneos de todas as conferências em andamento (nome da sala, quem está na

sala, pipelines, entre outros) ficam armazenados em um banco de dados distribuído a fim de aumentar a disponibilidade do sistema. Caso uma VM de GE fique indisponível, outra VM de GE assume, e utiliza os dados instantâneos que estão armazenados no BD distribuído, que é um elemento crítico do sistema para manter a alta disponibilidade.

- Monitor:** atua como um servidor que tem conhecimento do estado atual da nuvem de servidores de mídia. Seu papel é manter uma conexão com todos os agentes (ou clientes) monitores que estão presentes em cada uma das máquinas Servidores de Mídia (SMs) da nuvem, bem como receber, armazenar e informar o MCS sobre novos eventos detectados por esses agentes. Essa informação é fundamental para que o Balanceador de Carga (Gerenciador de Escalabilidade) consiga tomar uma decisão sobre qual Servidor de Mídia uma nova sala de videoconferência deve utilizar.

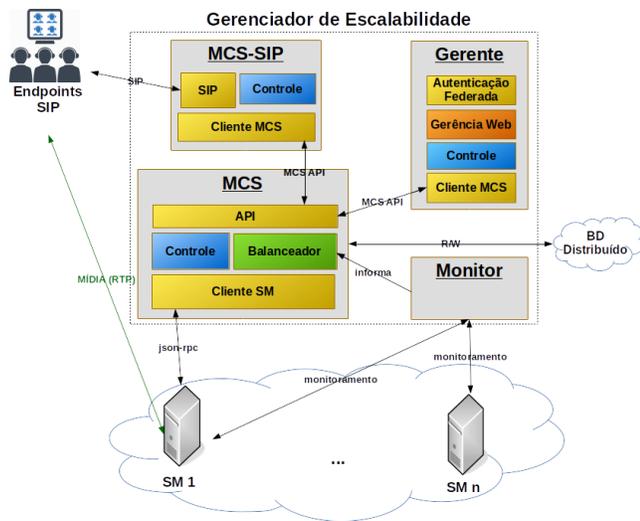


Figura 3: Esquema geral da arquitetura proposta

O padrão de sinalização proposto para o produto é o SIP (*Session Initiation Protocol*), permitindo a compatibilidade global do MCU proposto com os *endpoints* de mercado existentes.

Como pode-se perceber, a arquitetura proposta é facilmente escalável, totalmente baseada em máquinas virtuais na nuvem (tanto o GE e Banco de Dados como todos os SMs), permitindo futuras atualizações, troca de versões, aumento ou redução na capacidade do sistema, entre outras configurações. Isso torna o ambiente facilmente gerenciável caso seja oferecido como serviço por alguma NREN.

Adicionalmente à facilidade de gerência, a proposta de produto do GT-MCU é permitir acesso para a criação de salas virtuais de videoconferência de forma Federada. Dessa forma, a operação fica ainda mais facilitada, pois basta criar a regra e as próprias pessoas gerenciam suas salas.

A arquitetura proposta se destaca pois foi projetada de forma modularizada, com o intuito de facilitar a adição de novos componentes (por exemplo, módulos ou *plugins* do GStreamer,

ou novas aplicações desenvolvidas sobre o *framework*) que gerenciam e trabalham diretamente sobre as mídias (áudio, vídeo e conteúdo), independentemente do protocolo de sinalização das chamadas (SIP, H.323 ou novos que venham a surgir futuramente). Isso possibilita, por exemplo, que os vídeos de um ou mais usuários sejam interceptados e tratados antes e/ou após a sua composição, podendo se aplicar técnicas de visão computacional caso necessário (ex.: reconhecimento facial para identificação de pessoas, detecção de face para o posicionamento automático de câmeras frontais, etc).

Tecnologias Utilizadas

O sistema operacional utilizado na solução MCU é o Linux Ubuntu, que possui uma comunidade muito grande de desenvolvedores e dispensa maiores detalhes.

O principal componente do servidor de mídia é o software Kurento¹, que funciona através da licença LGPL v 2.1.

Para a decisão de utilização do Kurento como elemento principal do bloco Servidor de Mídia, o grupo comparou o mesmo com as seguintes soluções: a) FreeSwitch (<https://www.freeswitch.org/>); b) Licode (<http://lynckia.com/licode/>); c) Medooze (www.medooze.com/); e d) OpenMCU-ru (<https://videowswitch.ru/eng.htm>), tomando a decisão pelo mesmo em função da flexibilidade que a solução permite, habilitando tanto questões de MCU como SFU, e integrando facilmente com WebRTC.

O projeto Kurento existe desde 2012, surgindo dentro do Laboratório FUN-LAB (*Future Networks Laboratory*, ou Laboratório de Redes do Futuro) da Universidade Rey Juan Carlos (www.urjc.es) e com o objetivo de se tornar um servidor de mídia de código aberto com suporte a diferentes tipos de tecnologias e processamento de imagens.

Em termos de arquitetura, o Kurento é desenvolvido sobre a plataforma GStreamer (<https://gstreamer.freedesktop.org/>), que corresponde a um *framework* para desenvolvimento de aplicações multimídia, já utilizado em diversas aplicações Desktop e aplicações embarcadas, além de ser compatível com diferentes tipos de sistemas operacionais.

Outro software integrado no âmbito da proposta é o sistema de monitoramento Zabbix (www.zabbix.com/). O Zabbix é uma plataforma de código aberto para monitoramento escalável de infraestruturas de TI. Também realiza monitoramento baseado em agentes, podendo ser configurado como agente nas VMs dos Servidores de Mídia (SMs). Conta com uma interface web robusta, possui desenvolvimento e suporte ativo, sendo utilizado por diversas empresas consolidadas no mercado, como a Dell, Renner, ICANN, dentre outras.

O componente SIP utilizado para executar o padrão SIP é o SIP.js (<https://sipjs.com/>), que é uma biblioteca SIP em javascript também de código aberto. Essa biblioteca foi escolhida pois já possui várias sinalizações do protocolo SIP, evitando retrabalho, agilizando o desenvolvimento e garantindo mais compatibilidade.

¹ <https://www.kurento.org/>

4 TRABALHOS RELACIONADOS

Alonso [10] argumenta que um MCU é uma entidade que gerencia diferentes aspectos de sistemas multimídia, como a Mixagem, Encaminhamento, Gravação e Transcodificação de fluxos de mídia. Em seguida, apresenta cenários de uso de MCU em nuvem, e vantagens do modelo. Ele descreve os seguintes usos avançados de MCUs:

- **Streaming:** seria o uso do MCU para receber um fluxo de áudio e vídeo de uma fonte de origem (*publisher*) e , enviar o mesmo para vários assinantes.
- **Composição:** ao gerar um único fluxo de vídeo ou áudio a partir dos fluxos de entrada, o MCU reduz a sobrecarga nos *endpoints* e a banda na rede.
- **Transcodificação:** permite adaptar o fluxo de saída do MCU para diferentes necessidades de *endpoints*, como banda reduzida ou uma menor resolução, o que garante adaptação à heterogeneidade de dispositivos e redes de acesso.
- **Gravação:** permite manter o histórico das reuniões efetuadas.

Posteriormente, Alonso apresenta dificuldades de sistemas na nuvem, como a gerência da escalabilidade vertical (muitos numa sala) e horizontal (muitas salas). Ele apresenta também algumas vantagens dos sistemas em nuvem, como a distribuição geográfica, que permite o uso de uma nuvem próxima ao local da conferência sendo realizada.

Akkus [3] apresenta uma solução de multiponto que busca distribuir a carga do sistema utilizando *endpoints* como entidades capazes de gerir e processar localmente recursos de multimídia. Isso pode ser conseguido explorando peer-to-peer (P2P) ou estratégias de encaminhamento seletivo (SFU).

Kesavaraja [12] apresenta um conceito que está crescendo atualmente, que é o de oferecer serviços de videoconferência na nuvem (VaaS – *Video as a Service*).

Sistemas baseados em VaaS têm sido muito procurados para prover recursos de armazenamento, segurança e conferência. Armazenamento é a aplicação mais tradicional. A segurança desponta como preocupação recente, uma vez que provedores de serviços de multimídia têm sido alvos frequentes de hackers. Na videoconferência em nuvem, ainda são necessários software e hardware locais (*endpoints* providos de câmeras), mas os aplicativos de videoconferência que gerenciam as conferências e os dados que eles geram residem em algum ponto da rede (ou seja, na "nuvem"). Com isto se ganha em flexibilidade e mobilidade [13].

É importante destacar, entretanto, que aplicações de videoconferência baseados em VaaS, apesar de desejadas, ainda são emergentes, visto necessitarem gerência muito robusta para que não se mostrem sensíveis a flutuações e/ou congestionamentos na rede [14].

Rob Scott [15] apresenta um conceito de VGaaS (*Video Gateway as a Service*), onde a nuvem oferece um elemento central que converte protocolos entre si, permitindo a interação entre vários padrões, como SIP, WebRTC, ISDN, entre outros. Ele propõe o conceito de BYOC (*Bring Your Own Codec*), permitindo

unir numa mesma reunião várias pessoas sem se preocupar com qual equipamento, software ou codec cada um está utilizando.

Como pode ser visto através dos trabalhos relacionados, o conceito de utilização da nuvem como um elemento que efetua o papel de MCU ou de conversor de protocolos (gateway) está crescente atualmente.

5 RESULTADOS

Em termos de resultados, a ferramenta está funcional e sendo utilizada pelo grupo, em fase de início de projeto piloto com diferentes instituições. São efetuadas reuniões semanais do grupo de forma remota, através do MCU desenvolvido. Essa metodologia foi muito importante para encontrar erros no sistema, que foram corrigidos.

A Figura 4 mostra uma videoconferência efetuada pelo grupo, mostrando o layout com 5 participantes. Foram feitos experimentos com participantes em Brasília e até mesmo na Espanha, obtendo-se baixo atraso.



Figura 4: Videoconferência do grupo vista através do *endpoint* de software Jitsi

A Figura 5 mostra um resultado com o uso de um *endpoint* de hardware da Polycom. As quatro imagens representam o vídeo do Polycom (em cima à esquerda), um participante na mesma sala via Jitsi (embaixo à esquerda), e dois BOTs (à direita).



Figura 5: Videoconferência com um *endpoint* Polycom, um *endpoint* de software Jitsi e dois BOTS.

5.1 TESTES DE CARGA

Foram efetuados diversos testes de carga no sistema medindo CPU, memória, I/O, armazenamento, entre outros. Observou-se que o principal recurso computacional que oferece uma limitação ao número de usuários simultâneos é a CPU, visto a complexidade de codificação e decodificação de vídeo existente num MCU, conforme detalhado anteriormente. Fatores como memória e I/O não são afetados sensivelmente pelo processamento de vários vídeos.

É importante observar, também, que o ponto crítico de uso de CPU no sistema é nas máquinas virtuais (VMs) dos Servidores de Mídia (SM), e não do Gerenciador de Escalabilidade (GE). Isso se deve ao fato do GE trabalhar mais com sinalização (não onerando muitos recursos da máquina), e o SM trabalhar fortemente com codificação e decodificação de vídeo e áudio dos participantes das salas virtuais, onerando bastante os recursos computacionais das VMs.

Com base nisso, a Figura 6 apresenta o resultado obtido para o teste de carga específico de uso de CPU em uma sala virtual num dos servidores de mídia, tendo o número de vídeos sendo aumentado até 9. A partir desse ponto, a máquina somente recebia o vídeo do *endpoint*, mas o mesmo não entrava na composição. Obteve-se um limite aproximado de 25 usuários mantendo a qualidade inicial. Acima disso os vídeos começavam a diminuir o número de quadros por segundo (*frame rate*), diminuindo a qualidade da videoconferência.

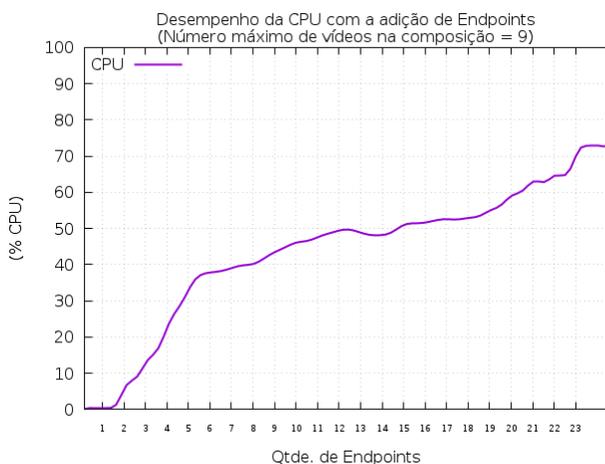


Figura 6: Resultado do teste de carga.

6 CONSIDERAÇÕES FINAIS

Este artigo apresentou um sistema de videoconferência do tipo MCU. Esse sistema já está em uso em forma de protótipo experimental na RNP. A arquitetura do MCU proposto é completamente inovadora em relação a MCUs tradicionais em hardware, possuindo funcionamento virtualizado em nuvem com possibilidade de escalabilidade. O fato de funcionar em ambiente virtualizado facilita muito a implantação e manutenção, diminuindo drasticamente os custos quando comparado a ambientes em hardware dedicado.

Algumas lições que podem ser extraídas do trabalho apresentado são:

- A arquitetura utilizada permite a escalabilidade em software de um sistema que antes do advento das máquinas virtuais seria muito difícil de implantar.
- O mecanismo de redundância proposto, onde os dados ficam num banco de dados redundante e distribuído, permite tolerância a falhas e aumenta a robustez da ferramenta.
- A limitação de uma sala do sistema é dada pela limitação de uma VM, o que deverá ser melhorado no futuro, fazendo com que o sistema permita uma sala suportada por múltiplas VMs.

Um vídeo que permite visualizar o sistema pode ser visto no início do texto em: <http://www.inf.ufrgs.br/prav/mcu.htm>.

AGRADECIMENTOS E LICENÇA

Este trabalho está sendo patrocinado pela RNP (Rede Nacional de Ensino e Pesquisa), a NREN brasileira, e pela Mconf Tecnologia. A licença de uso do sistema está para ser definida em conjunto com os parceiros envolvidos, entretanto, a base do sistema é o software Kurento e o Gstreamer, ambos abertos.

REFERÊNCIAS

- Kridel, Tim: The Value of Video Collaboration. Disponível em <https://www.avixa.org/insight/whitepapers/Details/the-value-of-video-collaboration/>. Novembro, 2017. Acesso em Maio, 2018.
- Sorokin, R.; Rougier, J-L. Video conference in the fog: an economical approach based on enterprise desktop grid. *Annals of Telecommunications*. 2017. pp. 1-12.
- Akkus, I. E. Civanlar, M. R. Ozkasap, O. "Peer-to-Peer Multipoint Video Conferencing using Layered Video", *Image Processing 2006 IEEE International Conference on*, Oct. 2006. pp. 3053-3056.
- <http://www.polycom.com.br/products-services/hd-telepresence-video-conferencing/realpresence-room/realpresence-room-hdx-series.html>. Acesso em Junho, 2018.
- <https://www.lifesize.com/en/video-conferencing-equipment>. Acesso em Junho, 2018.
- <https://www.cisco.com/c/en/us/products/collaboration-endpoints/telepresence-mx-series/datasheet-listing.html>. Acesso em Junho, 2018.
- <https://www.terena.org/activities/tf-webrtc/meeting1/slides/Jitsi.pdf>
- <https://www.ekiga.org/ekiga-softphone-features>
- <http://www.polycom.com/content/dam/polycom/common/documents/datasheets/realpresence-platform-virtual-editions-ds-enus.pdf>
- Alonso, A. et al. Deploying a Multipoint Control Unit in the Cloud: Opportunities and Challenges. In *Fourth International Conference on Cloud Computing, GRIDs, and Virtualization*. 2013 pp. 173-178.
- Roesler, V.; Longoni, G.; Marins, A. Multipresença: um sistema de videoconferência adaptável, escalável e interoperável. In: *TICAL 2015, 2015, Vina del Mar. Quinta Conferencia de Directores de Tecnologia de Información, TICAL 2015, 2015*.
- D. Kesavaraja and A. Shenbagavalli, "Cloud video as a Service [VaaS] with storage, streaming, security and Quality of service: Approaches and directions," *2013 International Conference on Circuits, Power and Computing Technologies (ICCPCT)*, Nagercoil, 2013, pp. 1093-1098.
- Liu W, et al. Cloud and traditional videoconferencing technology for telemedicine and distance learning. *Telem J E Health* 2015. n 21. pp. 422-426.
- Kang, J. A., Han, M. , Jang, J. and Kim, H. K. (2016), Adaptive Speech Streaming Based on Packet Loss Prediction Using Support Vector Machine for Software?Based Multipoint Control Unit over IP Networks. *ETRI Journal*, 38: 1064-1073.
- Scott, Rob: Video 2020 - The Future of Video Conferencing. Disponível em <https://www.uctoday.com/news/insights/video-2020-future-video-conferencing/>. Junho, 2017. Acesso em Maio, 2018.

Sistema Emagreça@Saudável para controle da obesidade em adultos

Maria Lúcia Kroeff Barbosa
PPGIE, UFRGS –
Universidade Federal do RGS
Porto Alegre, RS, Brasil
malukroeff@yahoo.com.br

Valter Roesler
Instituto de Informática, UFRGS –
Universidade Federal do RGS
Porto Alegre, RS, Brasil
roesler@inf.ufrgs.br

Luciano Zancan
Instituto de Informática, UFRGS –
Universidade Federal do RGS
Porto Alegre, RS, Brasil
lzancan@inf.ufrgs.br

Márcia Rosa da Costa
UFCSPA -
Universidade Federal de Ciências da Saúde
Porto Alegre, RS, Brasil
marciarc.ufcspa@gmail.com

Mário Figueiró Zemor
Instituto de Informática, UFRGS –
Universidade Federal do RGS
Porto Alegre, RS, Brasil
mario.ufrgs.inf@gmail.com

Sílvio César Cazella
PPGIE, UFRGS e UFCSPA –
Universidade Federal de Ciências da Saúde
Porto Alegre, RS, Brasil
silvioc@ufcspa.edu.br

ABSTRACT

This paper presents the Slim&Healthy system, focusing on the control of obesity in adults and aiming at the adoption of healthy habits. The system consists of two parts: a) an Android application on the client mobile device; b) an application on a server. The paper integrates different theories to increase user motivation, such as gamification, social networks, design heuristics for mobile devices and behavioral theories. It was performed an evaluation of the application interface through a questionnaire adherent to standards ISO 9241-11 and ABNT ISO/IEC 25062:2011, and the results are presented for discussion.

KEYWORDS

m-Health, obesity, health education

1 INTRODUÇÃO

Segundo um estudo publicado na revista *The Lancet* [1], o número de obesos no mundo cresceu seis vezes nas últimas quatro décadas, passando de 105 milhões de pessoas acima do peso em 1975 para 641 milhões em 2014. De acordo com as projeções, caso as pessoas continuem a engordar nesse ritmo, cerca de um quinto da população mundial estará acima do peso em menos de 10 anos, e o Brasil aparece como o 5º país mais obeso do mundo. Dados do Research2Guidance [2] apontam que existem 325.000 aplicativos de saúde nas principais lojas, e que os com maior potencial de sucesso para o m-Health estão relacionados ao tratamento de doenças crônicas como diabetes e obesidade, sendo que a América do Sul apresenta apenas 4% deste mercado. Desenvolver e validar um aplicativo que possa ser utilizado por uma grande parcela da população, com incentivos à mudanças de comportamento para prevenção e

tratamento da obesidade, mostra-se relevante à medida que visa apoiar ações de promoção e educação em saúde no Brasil.

Uma pesquisa anterior com as principais funcionalidades de aplicativos móveis direcionados ao controle da obesidade foi apresentada por Barbosa *et al.* [3], onde os cinco *apps* de maior destaque entre os dez selecionados foram analisados, a fim de identificar as funcionalidades indispensáveis. Com os resultados obtidos, foi possível aprimorar a modelagem do sistema.

O objetivo principal da ferramenta é fornecer uma solução *m-Health* para apoio à educação em saúde, com foco no controle de peso em adultos e visando a adoção de hábitos saudáveis, uma vez que a epidemia do sobrepeso e obesidade preocupa todas as nações, pois o excesso de peso é fator de risco elevado para doenças associadas à esta epidemia, doenças cardiovasculares (DCV), diabetes, doenças respiratórias e alguns cânceres [4].

2 EMAGREÇA@SAUDÁVEL

A Figura 1 apresenta o sistema Emagreça@Saudável e seus principais componentes e subcomponentes.

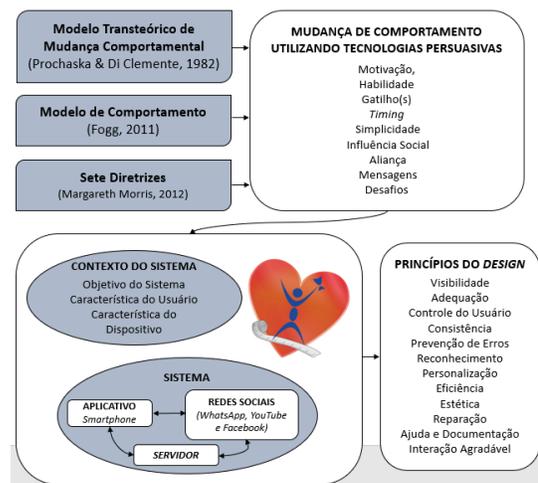


Figura 1: Visão global do sistema Emagreça@Saudável

Os três primeiros blocos azulados da Figura 1, Modelo Transteórico de Mudança de Comportamento [5], Modelo de Comportamento (*Fogg's Behaviour Model*¹) e Sete Diretrizes de Margaret Morris [6] ilustram as teorias comportamentais e diretrizes que embasam e apoiam a construção do sistema, relacionando-se para que ocorra uma efetiva mudança de comportamento com a utilização de tecnologias persuasivas, sendo esta uma abordagem reformulada do exposto por Barbosa *et al.* em outro artigo [7].

O bloco no quadrante inferior da direita ilustra as teorias utilizadas para a criação da interface com o usuário. Foram levados em consideração os princípios de *design* sugeridos por Inostroza *et al.* [8], voltados para interfaces de *smartphones* sensíveis ao toque com foco na usabilidade. O bloco no quadrante inferior da esquerda esquematiza o Emagreça@Saúdável, e seu contexto abrange:

- **Objetivo do Sistema:** disponibilizar um app para dispositivos móveis que envolva cuidados com a saúde, com incentivos a mudanças individuais de comportamento para prevenção e tratamento da obesidade, apoiando ações de educação em saúde.
- **Características do Usuário:** características gerais das pessoas que vão utilizar (idade, gênero, grau de escolaridade, etc.), bem como suas necessidades (estado de saúde), intenções e meio o qual se encontram.
- **Características do Dispositivo:** qualidades e limitações de uso de um *smartphone*.

Já o sistema (lado direito deste mesmo bloco) é composto por:

- **Aplicativo para dispositivos móveis:** o *app* é a interface principal, por onde o usuário cadastra seus dados, executa tarefas e desafios propostos, acompanha sua evolução e recebe orientações. Os dados do mesmo são armazenados localmente, e enviados ao Servidor.
- **Redes Sociais:** comunidade no Facebook, canal no *YouTube* e *WhatsApp* são usadas como apoio aos usuários do aplicativo, com repositório de vídeos, áudios e textos com dicas de saúde, esclarecimentos sobre estilo de vida, e sugestões de aulas diferenciadas. Também são espaços onde o usuário pode conversar e trocar experiências com outros usuários, bem como competir no modelo de gamificação. Pelo *WhatsApp* via *app*, é possível agrupar usuários de acordo com o perfil e enviar mensagens para grupos específicos. Tal funcionalidade é disponibilizada no Servidor.
- **Servidor:** possui o banco de dados que armazena todas as informações dos usuários, servindo como apoio para a comunicação com os dispositivos móveis e com as redes sociais. Trata-se de um sistema que, além da geração de estatísticas, permite o acompanhamento e atuação de uma rede multidisciplinar de profissionais. O conteúdo é inserido no Servidor pelos profissionais da área, e a arquitetura do sistema foi desenvolvida com esse foco de separação entre conteúdo e código. O Servidor também permite o envio de

mensagens direcionadas a certos perfis, bem como mensagens em geral. Isso é fundamental para a comunicação entre a equipe multidisciplinar e os participantes.

3 INTERFACE PROPOSTA

A seguir, a descrição de como esta interface foi realizada, tanto no dispositivo móvel, quanto no servidor.

3.1 No Dispositivo Móvel

Na primeira vez que o usuário entra no aplicativo, aparece uma mensagem de boas-vindas (Figura 2a) e a necessidade de fazer o *login* (ou *sign up*), permitindo inclusive que o usuário utilize seus dados do *Facebook* (Figura 2b).

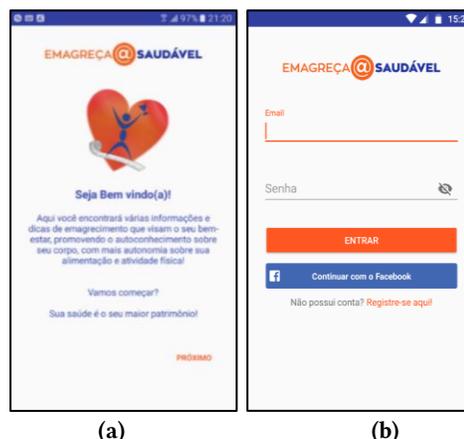


Figura 2: (a) Interface de boas-vindas; (b) Interface de login

Após, são solicitadas informações de perfil, tais como idade, escolaridade, uso de álcool/cigarro e medicamentos. Medidas de peso, altura (Figura 3a), cintura e quadril (3b) são essenciais, para calcular métricas de acompanhamento do usuário (IMC: Índice de Massa Corporal e ICQ: Índice Cintura Quadril).

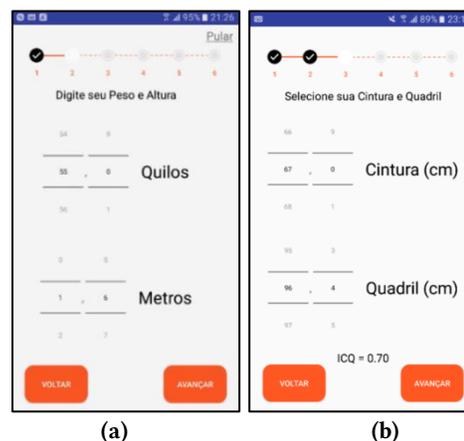


Figura 3: (a) Interface para entrada de Peso e Altura; (b) Interface para entrada de Cintura e Quadril

¹ <http://www.behaviormodel.org/>

A Figura 4(a) apresenta a mensagem de início da 1ª etapa (Onde Tudo Começa), procurando orientar o usuário neste primeiro momento. A Figura 4(b) mostra a tela “Tarefas”, com um “resumo” do andamento do usuário ao longo do seu processo.

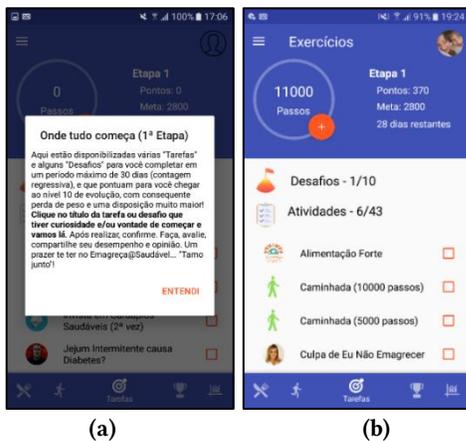


Figura 4: (a) Mensagem 1ª etapa; (b) Tarefas

Em destaque, na área superior de cor azul (Figura 4b), quantos passos o mesmo já realizou, qual sua pontuação e a meta de pontos a ser cumprida nesta etapa, além de quantos dias faltam para finalizar a mesma. Também apresenta opções de escolha de *Atividades* que devem ser cumpridas ao longo desta etapa. Quando as atividades são concluídas, aparece um ícone de verificação no quadrado ao lado do título da atividade e estas descem para o final da lista. As que ainda estão por fazer aparecem no topo da lista, bem como os *Desafios*, que entram como uma opção para o usuário concluir. As Atividades e Desafios foram pensados para que o usuário faça as tarefas solicitadas e entenda o motivo de sua realização, ou seja, possuem objetivos educacionais. Isso pode ser visto na Figura 5(a) nos itens “Cinco Dicas de Emagrecimento”, “Jejum Intermitente (Tudo Que Você Precisa Saber)”, “Por que você não faz o que sabe que deve fazer?”, entre outros que possibilitam ao usuário conhecer histórias de superação, dentro de um contexto parecido com o seu (Figura 5b).

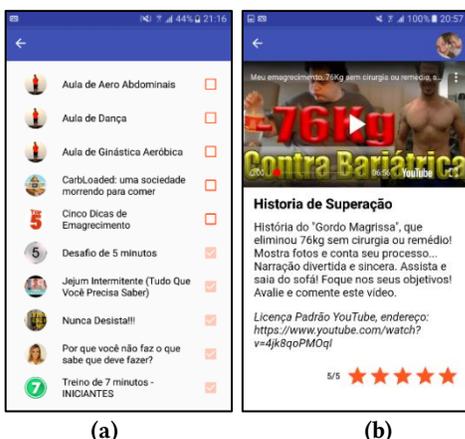


Figura 5: (a) Interface Desafios; (b) Interface com Exemplo de uma Atividade de Vídeo de Superação

A interface de abertura possui uma barra inferior azul (Figura 4) que permite o acesso rápido aos principais elementos da interface, além das tarefas, que são:

- **Alimentação:** a intenção não é que o usuário tenha que cadastrar toda refeição ingerida e ficar controlando calorias. A prioridade é trazer comparações e dicas de substituições de alimentos não saudáveis por mais saudáveis, bem como algumas sugestões e receitas, por exemplo, de águas saborizadas, alimentos saudáveis e cardápios funcionais. Além disso, também disponibiliza informações sobre jejum intermitente, entre outros.
- **Exercícios:** exibe uma lista de atividades físicas recomendadas com média de kcal gastas por tempo de treino. Tem opção de cadastro pelo usuário do tempo que realizou cada exercício para que contabilize em sua pontuação e o motive mais à prática de alguma atividade física.
- **Pontuação:** exibe a pontuação do usuário no sistema, buscando motivar o mesmo numa espécie de gamificação.
- **Evolução:** exibe gráficos de evolução de peso, IMC e ICQ do usuário, visando mostrar se houve ou não melhoria nas medidas e índices.

Exemplos destas interfaces são visualizados na Figura 6.



Figura 6: (a) Interface Alimentação; (b) Exercícios

Para quase todas as derivações apresentadas, o aplicativo traz em forma de texto, gráficos, áudios ou vídeos, informações e dicas que possam auxiliar o usuário a compreender as perdas de saúde que pode sofrer ao logo do tempo, caso opte por seguirem em uma rotina de autocuidados não satisfatória.

Na barra superior canto esquerdo da tela (Figura 6), aparece o ícone de . Clicando-se nele, é possível o acesso para:

- **Perfil:** aqui o usuário pode editar suas informações pessoais, bem a foto que deseja utilizar. Também visualiza suas medidas (iniciais e atuais), e tem a opção de ver as tabelas de IMC e ICQ, acompanhando seu progresso.

- **Motivacional:** o sistema envia mensagens automáticas. Algumas mensagens também são enviadas de forma direcionada, por meio da equipe interdisciplinar de apoio e em momentos adequados para cada um dos perfis de usuário.
- **Grupo WhatsApp:** integrar o participante em redes sociais a fim do mesmo trocar informações e mensagens com outras pessoas em situação similar, ou com interesse na troca de conhecimentos.
- **Configurações:** permite ao usuário ajustar as configurações do aplicativo, como definir horários a receber notificações de mensagens motivacionais e lembrete de medições semanais, por exemplo.
- **Perguntas Frequentes:** permite o acesso a perguntas mais frequentes sobre uso do aplicativo, bem como canais de atendimento disponibilizados via e-mail e WhatsApp (Ajuda).

A Figura 7 mostra as interfaces clicando-se no ícone  e item Perfil.

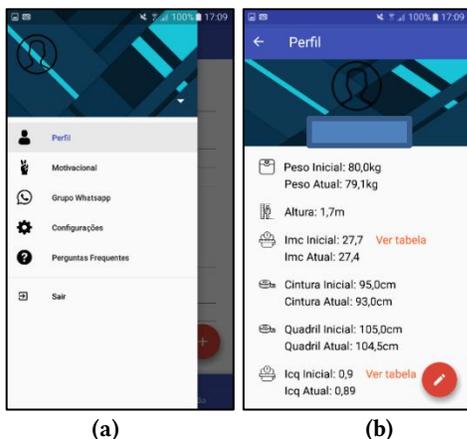


Figura 7: (a) Interface para o ícone ; (b) Interface Perfil

Foram apresentadas algumas das interfaces que compõem o aplicativo com o propósito de contemplar as recomendações para comunicação persuasiva com foco na mudança de comportamento para adoção de hábitos saudáveis. A seguir, será descrita a implementação no servidor.

3.2 No Servidor

A Figura 8 mostra um dos principais pontos que é o *Cadastro de Conteúdo*. O servidor foi idealizado para deixar a operação para uma equipe multidisciplinar sem maiores conhecimentos em programação.

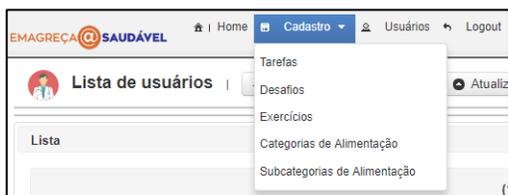


Figura 8: Cadastro de conteúdos

Sendo assim, o conteúdo é criado por especialistas nas áreas-chave do projeto, como Educação em Saúde, Nutrição, Psicologia, Educação Física, e inserido no banco de dados do Servidor, sendo automaticamente atualizado e baixado pelo aplicativo.

A Figura 9 mostra uma tela parcial da lista de conteúdos de *Tarefas* cadastrados no sistema. Pode-se observar que cada tarefa cadastrada pode ser editada, excluída e visualizada. Cada tarefa possui uma pontuação que determina o quanto vale a tarefa, na terceira coluna, e a classificação da tarefa (média da votação dos usuários, de 1 a 5 estrelas), na quarta coluna. As tarefas que estão com “0.0” não foram votadas ainda pelos participantes.

Sete Verdades sobre o Emagrecimento	vídeo	30	0,0	Visualizar	Editar	Excluir
Treinos Rápidos (1 minuto)	vídeo	10	5,0	Visualizar	Editar	Excluir
Treinos Rápidos (2 minutos)	vídeo	20	0,0	Visualizar	Editar	Excluir
Treinos Rápidos (3 minutos)	vídeo	30	0,0	Visualizar	Editar	Excluir
Treinos Rápidos (4 minutos)	vídeo	40	0,0	Visualizar	Editar	Excluir
Treino TABATA 1	vídeo	40	5,0	Visualizar	Editar	Excluir

Figura 9: Lista de tarefas cadastradas

Entrando em uma tarefa específica, é possível mudar o título, descrição, *link* da imagem ou vídeo, pontuação, entre outras, dependendo da tarefa, como pode ser visto na Figura 10.



Figura 10: Interface de visualização/edição de uma tarefa

O mesmo modelo vale para cadastro de Desafios, Exercícios, Categorias e Subcategorias de Alimentação. A lista de usuários permite visualizar rapidamente o perfil de cada participante (sexo, peso, IMC, ICQ), bem como sua pontuação e ranking (colocação) entre todos usuários. Ingressando em um usuário específico, é possível conhecê-lo melhor, saber o tipo de exercício que ele gosta mais, tarefas que realizou, dias na semana preferidos para se exercitar, bem como verificar graficamente a tendência de perda de peso, de medidas e de IMC e ICQ.

4 RESULTADOS

A avaliação da interface do *app* foi realizada com base nos requisitos de ergonomia fundamentados pela ISO 9241-11 [9] que trata a definição de usabilidade como: “medida na qual um produto pode ser usado por usuários especificados para alcançar objetivos especificados com eficácia, eficiência e satisfação em um contexto especificado de uso”. Um total de 35 questões procurou atender as 13 heurísticas a serem avaliadas, bem como

satisfação, utilidade e facilidade de uso, conforme sugere a norma Brasileira ABNT ISO/IEC 25062:2011 [10]. Utilizou-se uma escala de 1 a 5 (onde 1 = Discordo Totalmente, 2 = Discordo Parcialmente, 3 = Indiferente, 4 = Concordo Parcialmente e 5 = Concordo Totalmente). Além disso, caso o avaliador encontrasse problemas ou *bugs*, procurou-se buscar sugestões de correções e a prioridade de realização das mesmas.

Um total de 8 avaliadores especialistas (4 da área da Ciência da Computação, 2 da área da Informática na Educação, 1 da área da Psicologia e 1 da área da Saúde) utilizaram o *app* nos seus próprios *smartphones*. O período de avaliação foi de 20 de abril a 17 de maio de 2018. Conforme a norma ABNT ISO/IEC 25062:2011, 8 especialistas é um número suficiente para obter resultados significativos. A média dos resultados foram na grande maioria entre 4 e 5 (Concordo Parcialmente e Concordo Totalmente), em relação a contemplação de todos os aspectos avaliados. A Figura 11 apresenta o gráfico com a média dos resultados, bem como o intervalo mínimo e máximo das notas alcançadas, salientadas na linha de cor vermelha.

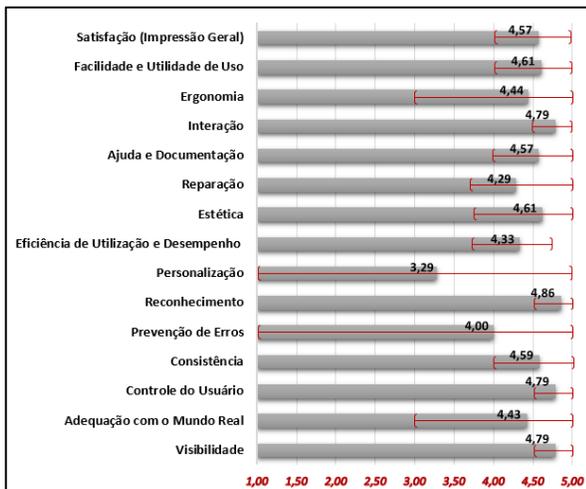


Figura 11: Resultados da avaliação da interface do aplicativo

Na “Personalização” (saber se o aplicativo mostrava opções de configurações), foi onde dois dos avaliadores sentiram dificuldade. O fato é que o aplicativo fornece alguns graus de personalização, permitindo ao usuário inserir seus dados, foto, escolher quem seguir (no processo de gamificação mencionado), além de prestar *feedbacks* de cada atividade efetuada e enviar mensagens ao grupo. Em “Prevenção de Erros”, somente um dos avaliadores avaliou na nota mínima em função de conseguir editar um número de passos muito acima do possível caminhar em um dia. Como o aplicativo não enviou uma mensagem de erro, solicitando que verificasse o valor digitado, ele considerou que tal heurística havia sido violada, com prioridade de correção. No mais, poucas dificuldades de uso foram encontradas, e já estão sendo melhoradas na nova versão do aplicativo.

5 CONSIDERAÇÕES FINAIS

O *m-Health* está em plena expansão e os aplicativos para saúde possuem grande potencial. Com as ferramentas necessárias para promoção da saúde e bem-estar, o sistema Emagreça@Saudável está embasado em teorias, modelos e diretrizes que possuem caráter persuasivo e visa auxiliar seus usuários não só na mudança, mas na permanência de um comportamento saudável. Além disto, reúne uma rede interdisciplinar de pesquisadores na área da obesidade, como psicólogos, nutricionistas, educadores físicos, médicos e programadores. Com tal rede e apoiado nos estudos efetuados, tal sistema visa contemplar diversos aspectos e procura manter o custo baixo de manutenção para difusão em larga escala. A ideia é manter o sistema livre, com política de uso de dados de forma anônima, mas abertos para novas pesquisas. O código-fonte do *app* não está concluído, mas a licença é aberta.

Como se pode observar, este trabalho permite também diversos desdobramentos, como a motivação dos usuários através da gamificação e o incentivo por meio das redes sociais. Pela quantidade potencial de dados dos participantes, vários estudos de *Data Mining* para *Big Data* podem ser efetuados com os dados obtidos pelo sistema. O *app* Emagreça@Saudável já se encontra disponível no Google Play para testes. Basta pesquisar na ferramenta de busca do Google Play, baixar, instalar e usar o mesmo.

REFERÊNCIAS

- [1] NCD Risk Factor Collaboration. 2016. Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *Lancet*. Vol 387 (Abril, 2016). DOI: [http://www.thelancet.com/pdfs/journals/lancet/PIIS0140-6736\(16\)30054-X.pdf](http://www.thelancet.com/pdfs/journals/lancet/PIIS0140-6736(16)30054-X.pdf)
- [2] RESEARCH2GUIDANCE. *mHealth App Developer Economics 2016: The current status and trends of the mHealth app market*. (Outubro, 2016). DOI: <http://research2guidance.com/r2g/r2g-mHealth-App-Developer-Economics-2016.pdf>
- [3] M. L. K. Barbosa *et al.* 2016. Aplicativos móveis para controle da obesidade e modelagem do emagreça@saudável. *RENOTE. Revista Novas Tecnologias na Educação*, v. 14, p. 1-10. DOI: <http://seer.ufrgs.br/renote/article/view/67371>
- [4] ABESO (Associação Brasileira para o Estudo da Obesidade e da Síndrome Metabólica). 2011. *Doenças Desencadeadas ou Agravadas pela Obesidade*. DOI: <http://www.abeso.org.br/uploads/downloads/28/5521afaf13cb9.pdf>
- [5] James O. Prochaska; Carlo C. DiClemente. 1982. *Transtheoretical therapy: Toward a more integrative model of change*. *Psychotherapy: Theory, Research and Practice*, v. 19, p. 276-288, 1982. DOI: <http://www.hbftpartnership.com/documents/uploadResources/Transtheoretic%20aT-Prochaska1982.pdf>
- [6] Margaret Morris. 2012. *Motivating change with mobile: seven guidelines*. *Interactions*, v.19, n. 3, p. 26-31. (Maio-junho, 2012). DOI: <http://pt.scribd.com/doc/92691136/Motivating-Change-With-Mobile-Seven-Guidelines-2012>
- [7] M. L. K. Barbosa *et al.* 2017. Educação permanente através de um sistema *m-Health* voltado ao controle da obesidade em adultos. *RENOTE. Revista Novas Tecnologias da Educação*, v. 15, p. 1-10. DOI: <http://seer.ufrgs.br/renote/article/view/75113/42553>
- [8] Rodolfo Inostroza *et al.* 2013. *Usability heuristics for touchscreen-based mobile devices: Update*. DOI: <http://jcc2013.inf.ucl.cl/wp-content/proceedings/ChileCHI/Usability%20Heuristics%20for%20Touchscreen-based%20Mobile%20Devices%20Update.pdf>
- [9] OBP (*Online Browsing Plataforma*). ISO 9142-11:2018 (en). *Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts*. DOI: <https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en>.
- [10] NBRISO/IEC25062: Engenharia de software - Requisitos e avaliação da qualidade de produto de software (SQuaRE). 2011. DOI: <https://www.target.com.br/produtos/normas-tecnicas/42260/nbriso-iec25062-engenharia-de-software-requisitos-e-avaliacao-da-qualidade-de-produto-de-software-square-formato-comum-da-industria-fci-para-relatorios-de-teste-de-usabilidade>.

Digital Cognitive Tests with Brazilian elderly

Leandro Agostini do Amaral
Universidade de São Paulo
Avenida Trabalhador São-carlense,
400 – Centro
São Carlos, São Paulo 13566-590
leandroagostini@usp.br

Renata Pontin de Mattos Fortes
Universidade de São Paulo
Avenida Trabalhador São-carlense,
400 – Centro
São Carlos, São Paulo 13566-590
renata@icmc.usp.br

Thiago Jabur Bittar
Universidade Federal de Goiás
Av. Dr. Lamartine Pinto de Avelar,
1120
Catalão, Goiás 75704-020
thiagojabur@gmail.com

ABSTRACT

Serious games are presented as a very promising mechanism in many settings, such as education and health, however its application with the elderly is little verified. In order to reproduce in digital form the results of the elderly cognitive assessments, made and validated on paper, in this study a digital cognitive test tool, with the characteristics of a serious game, was developed using the Unified Design process. As conclusion, it is suggested that it is possible to achieve results similar to those obtained by validated and paper-based tests, from a game, adapted to the needs of the elderly, in a digital cognitive test format.

KEYWORDS

Digital Cognitive Test, Serious Games, Game Accessibility

1 INTRODUÇÃO

O envelhecimento populacional é uma realidade mundial e no Brasil não é diferente. Segundo projeção realizada pelo sistema do IBGE [1], em novembro de 2017, a população de 50 anos, ou mais, integrava 23,34% da população brasileira, o que representa aproximadamente 48,5 milhões de habitantes. Esse percentual tende a aumentar, com previsões de que em 2030, alcance 31,2%. Um grande desafio enfrentado com o envelhecimento populacional é a maior prevalência de transtornos neurocognitivos. Essa prevalência na população mundial acima de 65 anos varia entre 5 e 7%, com maiores valores (8,5%) encontrados na América Latina [2]. No Brasil, os estudos de prevalência encontraram taxas que variaram de 5,1% [3] a 12,9% [4, 5]. Em número absoluto, estima-se que o Brasil seja o nono país do mundo com maior número de casos de demência, cerca de um milhão de indivíduos [2].

As demências constituem uma questão de saúde pública e trazem uma preocupação bioética: a perda da autonomia dos afetados e a responsabilidade do sistema de saúde para atendê-los [4].

O sistema biológico mais comprometido com o processo de envelhecimento é o Sistema Nervoso Central (SNC), responsável por gerar as sensações, os movimentos, funções psíquicas e

funções biológicas internas [6]. A integridade da estrutura dos neurônios e das vias neuronais torna o SNC apto ao acúmulo de informações do presente, lembranças do passado e à formulação de novos conceitos. Assim como outros sistemas biológicos, o SNC também possui propriedades capazes de diminuir os impactos causados pelas alterações do envelhecimento [6].

O próprio envelhecimento saudável acarreta, aos idosos, alterações cognitivas [7]. As alterações mais comuns ocorrem na função executiva, atenção, memória episódica, memória de trabalho e percepção [7, 8]. A função executiva é um conjunto de habilidades, que de forma integrada, possibilita ao indivíduo direcionar comportamentos a objetivos, realizando ações voluntárias. Tais ações são auto organizadas, mediante a avaliação de sua adequação e eficiência em relação ao objetivo pretendido, de modo a eger as estratégias mais eficientes, resolvendo assim, problemas imediatos, e/ou de médio e longo prazo. A função executiva é requerida sempre que se faz necessário formular planos de ação ou quando uma sequência de respostas apropriadas deve ser selecionada [9]; atenção é a tomada de posse, pela mente, de modo claro e vívido, de um entre os muitos objetos ou cadeias de pensamento possíveis. A focalização e a concentração da consciência fazem parte da sua essência. Ela implica o afastamento de algumas coisas de modo a que se possa lidar efetivamente com outras [10]; memória episódica é um sistema neurocognitivo que possibilita ao ser humano recordar fatos passados, podendo realizar uma "viagem no tempo", relembrando episódios de sua vida até o presente instante [11]; memória de trabalho representa um sistema de memória ultrarrápida (dura poucos segundos), que tem a capacidade de reter, por exemplo, uma sequência de 5 a 9 dígitos - o suficiente para gravar um número de telefone até efetuar a discagem, esquecendo-o logo em seguida [9] e percepção é o processo pelo qual o cérebro seleciona, organiza e interpreta estímulos, traduzindo-os em uma imagem significativa [12].

O declínio cognitivo pode comprometer a saúde do idoso, com consequente perda da capacidade funcional e autonomia, resultando em diminuição da qualidade de vida [13]. A cognição é cada vez mais reconhecida como um fator determinante para a resiliência e a homeostase corporal [14].

Como alternativa para retardar os danos à função cognitiva inerentes e que naturalmente ocorrem devido ao processo de envelhecimento, pesquisas têm sido realizadas sobre os efeitos benéficos dos jogos digitais, os quais podem revolucionar a avaliação cognitiva dos idosos, em ambientes clínicos, permitindo

In: XVII Workshop de Ferramentas e Aplicações (WFA 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pôsteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.
©2018 SBC – Sociedade Brasileira de Computação.
ISBN: 978-85-7669-435-9.

que as avaliações sejam mais frequentes, mais acessíveis e mais agradáveis [15].

Foram investigadas as principais ferramentas para teste cognitivo, atualmente existentes, que já são validadas para uso em papel. A seguir, foi proposta uma conversão de teste cognitivo para uma ferramenta em modelo digital, seguindo o processo estruturado de design de jogos, denominado Design Unificado [16].

2 TESTE COGNITIVO DIGITAL PARA IDOSOS – CONCEITOS

Os trabalhos estudados, com base em pesquisas publicadas e relacionadas com os tópicos desta pesquisa, possibilitaram o projeto e desenvolvimento de uma ferramenta de teste cognitivo digital voltado ao público idoso, que atendesse à demanda de proporcionar similaridade com a atividade de avaliação por meio de baterias de teste cognitivo, comumente utilizadas por profissionais da saúde.

Nesse contexto, são apresentados os conceitos de *serious games* e seus benefícios, da avaliação cognitiva e os domínios cognitivos, e de um conjunto de técnicas para reduzir as barreiras ao ambiente de jogos digitais, especialmente relativas à acessibilidade, pelo público idoso.

2.1 Serious Games

Diante dos vários domínios de conhecimento que já abordaram o tema de jogos, como: educação [17,18], saúde [19], religião [20] e política [21], e atendo-se ao segmento que detém o entretenimento como segundo plano, emerge a denominação *serious games*.

De modo geral, a expressão “*serious games*” é adotada com o significado de que os jogos dessa categoria possuem uma motivação educativa explícita, cuidadosamente pensada, e não se destinam a serem jogados exclusivamente para diversão ou entretenimento. No entanto, essa interpretação não significa que *serious games* não sejam, ou não devam ser, divertidos; ao contrário, a ideia é de que *serious games* também possuam elementos criativos e agradáveis como seus componentes.

Os *serious games* são projetados para terem um impacto no público-alvo, que está além do entretenimento somente [18, 22, 23]. Um dos domínios de aplicação mais importantes é o da Educação, dado seu potencial reconhecido frente à contínua necessidade de aprimoramento das técnicas e atualização de recursos educacionais [24].

Outra iniciativa mais recente para uso dos *serious games* é sua adoção pelo público idoso. Esse público-alvo representa uma comunidade de potenciais usuários que podem se beneficiar muito de jogos digitais. Argumenta-se que por meio de jogos digitais, é possível amenizar o declínio de alguns dos aspectos inerentes ao envelhecimento [25].

Em uma pesquisa sobre diferentes momentos e situações clínicas dos idosos, os estudos de [26] revelaram que *serious games* têm se mostrado muito importantes para manter e desenvolver as competências cognitivas e sociais dos idosos. Na literatura especializada, estudos apontam que *serious games* para idosos podem ser uma opção para melhorar os danos das funções

cognitivas [27] e motoras [28], a manutenção da autoimagem e o convívio social (afetividade e solidão) [29].

Segundo [30], uma área bastante promissora para aplicação de *serious games*, voltados para o público idoso, é a avaliação neuropsicológica, que consiste em uma investigação minuciosa das funções cognitivas, sensoriais, motoras, emocionais e sociais da pessoa, com o intuito de se identificar algum comprometimento funcional neurológico, bem como as áreas preservadas. Podem também proporcionar outros benefícios, inerentes aos testes cognitivos computacionais [31]: melhorar a padronização da administração de testes; prover fontes mais ricas de dados sobre medidas de desempenho, como tempo de resposta, latência e variabilidade; capacitar a avaliação de um maior número de domínios; aumentar a sensibilidade; reduzir vieses culturais; aumentar a consistência entre os grupos clínicos; aumentar a eficiência e controle de qualidade para o gerenciamento de dados; ser menos estressante e fornecer maior senso de domínio e controle aos participantes.

Com base nesses trabalhos, a adoção de *serious games* nesta pesquisa se justifica. Na próxima subseção, são descritos os trabalhos que foram estudados, sobre a atividade de avaliação cognitiva, que tem sido realizada por profissionais da área de Saúde dos idosos.

2.2 Avaliação Cognitiva

Na avaliação cognitiva global, a grande maioria dos estudos utilizam escalas de rastreamento breves, como por exemplo o Mini Exame do Estado Mental (MEEM) [14]. O MEEM é um instrumento de rastreamento (verificação) amplamente utilizado, que apresenta escore que varia de 0 a 30, e avalia orientação temporal e espacial, memória (fixação e evocação), linguagem, atenção e cálculo. Diversos estudos de validação para determinar suas qualidades psicométricas foram realizados no Brasil, em diferentes populações [32].

Outros instrumentos para avaliação de desempenho cognitivo, que foram reportados na literatura, especialmente para uso da população brasileira, são os seguintes:

1. *Lista de figuras da Bateria breve de rastreio cognitivo (BBRC)*: utilizando-se de figuras para avaliação de memória. Este teste consiste nas seguintes etapas: nomeação, memória incidental, memória imediata, aprendizagem, memória tardia e reconhecimento [33];
2. *Teste de extensão de dígitos (ordem inversa e direta) da Wechsler Memory Scale revised*: constituído por sete pares de sequências numéricas, com diferentes quantidades de dígitos, o teste é aplicado na sequência direta e inversa, sendo que durante a aplicação direta, as sequências têm de 3 a 9 números e na inversa de 2 a 8 números [34]. O teste termina após erro de duas sequências consecutivas, sendo registrada a quantidade máxima de números repetidos sem erro para cada versão;
3. *Teste de nomeação de Boston*: é o teste para avaliar linguagem mais utilizado no Brasil, principalmente para avaliar a capacidade de nomeação. É composto por 15 figuras, com três graus de dificuldade. A aplicação consiste em solicitar ao participante nomear os desenhos. O teste apresenta valores normativos para o Brasil [35];

4. *Teste do Desenho do Relógio - TDR*: é comumente empregado na investigação da presença de comprometimento de algumas habilidades cognitivas, como funções visuoespaciais e visuoespaciais, representação simbólica e grafomotora, memória semântica e funções executivas. A pontuação consiste em uma escala que varia de 0 a 10 pontos. O teste encontra-se validado na população brasileira [36];
5. *Teste de fluência verbal semântica*: avalia diferentes domínios cognitivos como linguagem, memória semântica e função executiva. O examinador pede ao sujeito que fale o maior número possível de animais, no período de um minuto. A avaliação dependerá do grau de escolaridade do participante: a nota de corte para indivíduos com até oito anos de escolaridade é 9, e para indivíduos com mais de 8 anos de escolaridade, a nota de corte é 13 [37];
6. *Lista de palavras do CERAD*: uma lista de palavras é apresentada para o sujeito e ao final de cada apresentação deve-se recordar o maior número de palavras possível. O procedimento é repetido por três vezes. Após um período de tempo, o examinado será solicitado a evocar as palavras e em seguida reconhecer as mesmas em uma lista de 20 palavras [38];
7. *Figuras geométricas do CERAD*: é um subteste da bateria do CERAD [38], no qual é solicitado ao paciente desenhar uma cópia de quatro figuras geométricas simples;
8. *Subtestes de semelhanças CAMDEX*: constituído de quatro perguntas para avaliar a capacidade de abstração dos participantes, por exemplo, o que tem em comum uma maçã e uma banana [39];
9. *Reconhecimento de emoções faciais - REF*: pode ser feito pelo *Penn Emotion Recognition Test*, no qual fotografias de expressões faciais são apresentadas, com as seguintes emoções básicas: alegria, tristeza, raiva, medo e nojo, e faces/rostos sem emoção (neutro).

A categorização dos instrumentos estudados, separados por domínios neurocognitivos também deve ser levada em consideração de acordo com o *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)* [40].

2.3 Acessibilidade no Âmbito de Jogos Digitais

Apesar do interesse crescente pelo uso de jogos digitais em diferentes áreas, muitas pessoas, dentre as quais os idosos, acabam por serem privadas de utilizá-los, seja por algum declínio, restrição visual, motora, auditiva ou cognitiva [41–43].

Para considerar as limitações dos indivíduos que podem ser atingidos pela falta de acessibilidade, utiliza-se, em geral, a classificação definida pela *World Health Organization's (WHO)* em *International Classification of Impairments, Disabilities, and Handicaps (ICIDH)*, que abrange os seguintes comprometimentos: *visual*, que é a consequência de um certo grau de perda de visão; *auditivo*, o qual refere-se à perda parcial ou total da habilidade de ouvir, por um ou pelos dois ouvidos; *motor*, que é a perda ou limitação da função de controle muscular, de movimento ou limitação da mobilidade; *cognitivo*, que representa uma enfermidade mental ou psicológica, que vai desde um retardo desenvolvido durante a infância, até Alzheimer ou senilidade, como resultado do processo de envelhecimento.

Conforme a popularidade da tecnologia aumenta, intensificam-se os esforços para compreender aqueles atingidos por esses

comprometimentos, bem como tornar a tecnologia acessível ao maior número de pessoas. Um resultado desse direcionamento é referência mundial e consiste na *Web Accessibility Initiative (WAI)*, que elaborou o documento *Web Content Accessibility Guidelines (WCAG)*, desenvolvido a partir do processo [44] do *World Wide Web Consortium (W3C)*;

As diretrizes WCAG orientam a produção de conteúdo (textos, imagens, formulários, sons) para a web, que seja acessível às pessoas com a diversidade de tipos de deficiência.

Com a realidade do aumento da popularidade dos jogos digitais, o enfoque de trabalhos e pesquisas no sentido de prover acessibilidade a esse tipo de sistemas interativos também deve aumentar, porém ainda são poucos os esforços neste sentido [45]. Não há diretrizes oficiais, padrões ou iniciativas globais, comparáveis às WCAG, no domínio de jogos digitais, tampouco ações governamentais ou legislativas relacionadas, entretanto, existem duas intervenções no sentido de obter esse conjunto de diretrizes: uma publicada pelo *Special Interest Group (SIG)*, da *Independent Game Developers Association (IGDA)*, que propôs 19 diretrizes de acessibilidade, em 2004, e atualizadas em 2010, obtidas de um experimento com 20 jogos, priorizando os usuários com deficiência visual e outra, da organização norueguesa *MediaLT*, que também em 2004, publicou um conjunto de 34 diretrizes de acessibilidade.

Mais recentemente, em 2012, visando produzir uma referência objetiva aos desenvolvedores de jogos digitais, e incluir o maior número de pessoas possível, considerando as diferentes necessidades apresentadas por pessoas com deficiências, um esforço colaborativo entre produtores, especialistas e acadêmicos foi realizado, surgindo as diretrizes de acessibilidade para jogos.

Paralelamente à obtenção de diretrizes de acessibilidade para jogos digitais, formalizadas internacionalmente, evidencia-se o avanço da qualidade dos jogos, a partir da criação de um processo estruturado de design de jogos, denominado *Design Unificado (DU)* [16]. O DU inclui como uma de suas etapas, a avaliação de acessibilidade técnica, que pode verificar o uso com base nas diretrizes, mas também recorrer à participação de especialistas e usuários finais.

3 MÉTODO E RESULTADOS

Este estudo é de natureza transversal de levantamento de dados, descritivo e analítico, a partir de uma investigação qualitativa. A pesquisa foi realizada na cidade de São Carlos, situada no Estado de São Paulo, Brasil.

Para o desenvolvimento do teste cognitivo digital, considerou-se o processo DU na fase de análise e o *Integrated Development Environment (IDE) Stencyl* foi utilizado para desenvolvimento.

O tablet foi escolhido como dispositivo final, considerando a eficácia e assertividade no uso pela população idosa, que, com dificuldades psicomotoras e audiovisuais, encontra, nesse dispositivo, uma ferramenta mais acessível ao ambiente virtual, vencendo algumas barreiras relativas às dificuldades psicomotoras.

O estudo, para identificar e projetar a interface mais adequada ao público idoso, considerou questões de acessibilidade desde seu planejamento, por meio do processo DU e pesquisas sobre como avaliar a acessibilidade em jogos digitais, especialmente para esse público.

Um exemplo de decisão tomada durante o desenvolvimento do protótipo, além da observação das diretrizes de acessibilidade, foi

a escolha da paleta de cores, realizada respeitando-se a possibilidade de atender indivíduos que apresentassem daltonismo, mais especificamente Tritanopia, Protanopia e Deuteranopia. A ferramenta Color Oracle foi utilizada para averiguar a paleta nas três situações e então foram eliminadas cores que se aproximavam umas das outras, tornando-as quase indistinguíveis. Ainda assim foi necessário o devido cuidado ao confeccionar os *assets* do jogo, isto é, *feedbacks* e objetivos não podem contar apenas com a cor para diferenciar objetos; portanto, sons e formas também foram utilizados.

Na bateria de testes cognitivos, foram apurados nove aspectos, a saber: 01 - Percepção Visual, Memória Incidental e Memória Imediata, 02 - Praxia, 03 - Abstração, 04 - Percepção Visual, Memória Incidental e Memória Imediata, 05 - Teste de Extensão de Dígitos (Direto), 06 - Teste de Extensão de Dígitos (Inverso), 07 - Atenção, 08 - Teste de Fluência Verbal e 09 - Reconhecimento Facial. Esses aspectos foram apurados e serviram para analisar uma possível correlação, que demonstrasse a portabilidade da avaliação convencional (em papel), para o meio digital, sem que houvesse uma alteração significativa no contexto. Nas Figuras 1, 2, 3, apresentadas a seguir, é possível verificar algumas das telas da ferramenta, com elementos de fácil utilização pelo tamanho e cores para a execução dos testes cognitivos.

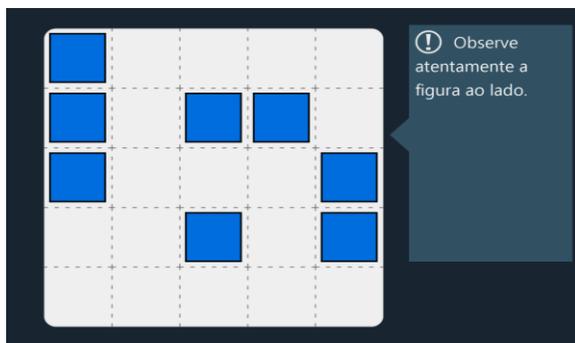
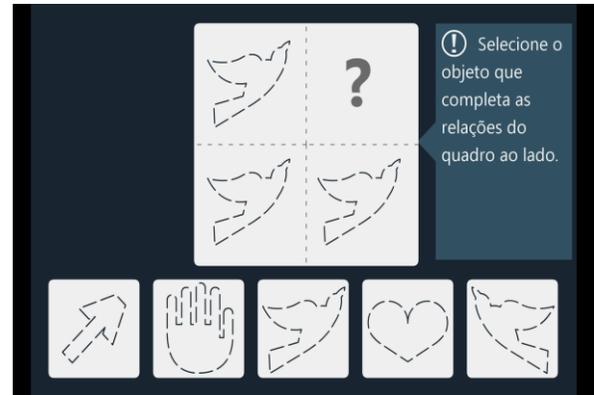


Figura 1: Tela para obter instruções sobre como identificar a diferença na matriz.



Figura 2: Tela para relatar falhas ao lembrar uma sequência de números.



Teste baseado no Beta III.

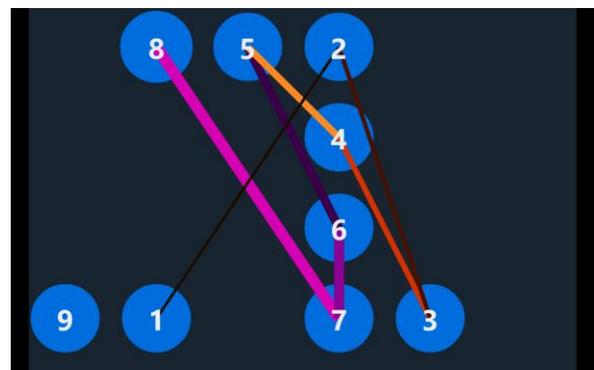


Figura 3: Teste baseado no Trail Making Test.

Para cada um dos testes, tem-se medidas (atribuídas a variáveis), que foram capturadas durante o processo de interação do idoso. Essas medidas foram utilizadas na fase de análise da convergência, com os testes validados em papel. O objetivo principal no uso dessas variáveis é representar o resultado de cada teste e possibilitar análise computacional, sem a percepção humana, a fim de inibir vieses por parte do aplicador. Um vídeo demonstrando a execução do teste cognitivo digital, resultado desse estudo, pode ser visto em: <http://leandroagt.com.br/video.mp4>.

5 CONCLUSÕES

Os *serious games* têm se mostrado relevantes quando utilizados em diversos tipos de aplicações, como é o caso de apoiar atividades no campo da saúde. Nesta pesquisa, foi investigada sua adoção considerando procedimentos de avaliação cognitiva com o público idoso do Brasil. De fato, as tecnologias digitais se tornam cada vez mais uma realidade para o mundo atual, e os novos idosos já têm sido introduzidos a elas, muitas vezes, desde seu próprio ambiente familiar ou de trabalho.

Assim, esta ferramenta que interliga o potencial de jogos, por meio de *serious games*, à demanda da sociedade atual, cujo contingente populacional se expande rapidamente, se apresenta como um quesito mais que natural. Esse contexto motivou a pesquisa descrita, a qual sugere que é possível atingir resultados

semelhantes, a partir de um jogo, adaptado às necessidades do idoso, em formato de teste cognitivo digital, aos obtidos pelos testes baseados em papel, já validados e comumente utilizados por profissionais da saúde. Com esse resultado, poder-se-á utilizar o teste cognitivo digital para o âmbito de pesquisa acadêmica, bem como na identificação de perfis dos usuários para um projeto de produto em andamento, que é o ambiente de treinamento cognitivo, por meio de jogos digitais, direcionados aos declínios identificados no jogador.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001 e pela FAPESP - projeto número 2016/01009-0.

REFERÊNCIAS

- [1] IBGE. Projeção da população, <http://www.ibge.gov.br/apps/populacao/projecao/>
- [2] Prince M., Bryce R., Albanese E., Wimo A., Ribeiro W., Ferri C.P.: The global prevalence of dementia: a systematic review and meta-analysis *Alzheimers Dement.*, 9, pp. 63–75. e2 (2013)
- [3] Scazufca M., Menezes P.R., Vallada H.P., Crepaldi A.L., Pastor-Valero M., Coutinho L.M.S., Di Rienzo V.D., Almeida O.P.: High prevalence of dementia among older adults from poor socioeconomic backgrounds in São Paulo, Brazil *Int. Psychogeriatr.*, 20, pp. 394–405 (2008)
- [4] Burlá C., Camarano A.A., Kanso S., Fernandes D., Nunes R.: Panorama prospectivo das demências no Brasil: um enfoque demográfico *Cien. Saude Colet.*, 18, pp. 2949–2956 (2013)
- [5] Bottino C.M.C., Azevedo D., Jr., Tatsch M., Hototian S.R., Moscoso M.A., Folquitto J., Scalco A.Z., Bazzarella M.C., Lopes M.A., Litvock J.: Estimate of Dementia Prevalence in a Community Sample from São Paulo, Brazil *Dement. Geriatr. Cogn. Disord.*, 26, pp. 291–299 (2008)
- [6] Py L., de Freitas E.V., do Gorzoni M.L.: Tratado de geriatria e gerontologia, Grupo Gen - Guanabara Koogan, (2011)
- [7] Ferreira D., Molina Y., Machado A., Westman E., Wahlund L.-O., Nieto A., Correia R., Junqué C., Díaz-Flores L., Barroso J.: Cognitive decline is mediated by gray matter changes during middle age *Neurobiol. Aging*, 35, pp. 1086–1094 (2014)
- [8] Mishra J., Rolle C., Gazzaley A.: Neural plasticity underlying visual perceptual learning in aging *Brain Res.*, 1612, pp. 140–151 (2015)
- [9] Mourão Junior C.A., Junior C.A.M., Melo L.B.R.: Integração de três conceitos: função executiva, memória de trabalho e aprendizado *Psicologia: Teoria e Pesquisa*, 27, pp. 309–314 (2011)
- [10] Kandel E.R., Rubino R.: Em busca da memória: o nascimento de uma nova ciência da mente, (2009)
- [11] Tulving E.: Episodic Memory: From Mind to Brain *Annu. Rev. Psychol.*, 53, pp. 1–25 (2002)
- [12] Lamb C.W., Hair J.F., McDaniel C.: Marketing, South Western Educational Publishing, (2004)
- [13] Grden C.R.B., Barreto M.F.C., Sousa J.A.V. de, Chuertniek J.A., Reche P.M., Borges P.K. de O.: Association between physical frailty and cognitive scores in older adults *Rev Rene*, 16, pp. 391–397 (2015)
- [14] Canevelli M., Cesari M., van Kan G.A.: Frailty and cognitive decline: how do they relate? *Curr. Opin. Clin. Nutr. Metab. Care*, 18, pp. 43–50 (2015)
- [15] Tong T., Chignell M., Tierney M.C., Lee J.: A Serious Game for Clinical Assessment of Cognitive Status: Validation Study *JMIR Serious Games*, 4, pp. e7 (2016)
- [16] Grammenos D., Savidis A., Stephanidis C.: Designing Universally Accessible Games *Human Factors and Ergonomics*, pp. 1–12 (2009)
- [17] Gee J.P.: What video games have to teach us about learning and literacy *Computers in Entertainment*, 1, pp. 20 (2003)
- [18] Kebritchi M., Hirumi A.: Examining the pedagogical foundations of modern educational computer games *Comput. Educ.*, 51, pp. 1729–1743 (2008)
- [19] Lieberman D.A.: Interactive Video Games for Health Promotion: Effects on Knowledge, Self-Efficacy, Social Support, and Health Health Promotion and Interactive Technology: Theoretical Applications and Future Directions. pp. 103–120 (1997)
- [20] NBC. Give me that online religion, http://www.nbcnews.com/id/18789168/ns/nbc_nightly_news_with_brian_williams/t/give-me-online-religion/
- [21] Delhi J.P. in: Tiny island nation opens the first real embassy in virtual world, <https://www.thetimes.co.uk/article/tiny-island-nation-opens-the-first-real-embassy-in-virtual-world-c3phx2n2r0>
- [22] Michael D.R., Chen S.: Serious Games: Games that Educate, Train and Inform, Course Technology, (2006)
- [23] Greitzer F.L., Kuchar O.A., Huston K.: Cognitive science implications for enhancing training effectiveness in a serious gaming context. *ACM J. Educ. Resour. Comput.*, 7, pp. 2–es (2007)
- [24] Frederik D.G., Peter M., Van Looy J.: Uncharted waters? *Proceedings of the 3rd International Conference on Fun and Games - Fun and Games '10* (2010)
- [25] Chen S.-T., Huang Y.-G.L., Chiang I.-T.: Using Somatosensory Video Games to Promote Quality of Life for the Elderly with Disabilities 2012 *IEEE Fourth International Conference On Digital Game And Intelligent Toy Enhanced Learning* (2012)
- [26] Programa de estimulação cognitiva em idosos institucionalizados, http://www.psicologia.pt/artigos/ver_artigo.php?codigo=a0623
- [27] Torres A.C.S.: Cognitive effects of video games on old people *Int. J. Disabil. Hum. Dev.*, 10, (2011)
- [28] Wiemeyer J., Kliem A.: Serious games in prevention and rehabilitation—a new panacea for elderly people? *Eur. Rev. Aging Phys. Act.*, 9, pp. 41–50 (2011)
- [29] Jung Y., Li K.J., Janissa N.S., Gladys W.L.C., Lee K.M.: Games for a better life *Proceedings of the Sixth Australasian Conference on Interactive Entertainment - IE '09* (2009)
- [30] Valladares-Rodríguez S., Pérez-Rodríguez R., Anido-Rifón L., Fernández-Iglesias M.: Trends on the application of serious games to neuropsychological evaluation: A scoping review *J. Biomed. Inform.*, 64, pp. 296–319 (2016)
- [31] Espeland M.A., Katula J.A., Rushing J., Kramer A.F., Jennings J.M., Sink K.M., Nadkarni N.K., Reid K.F., Castro C.M., Church T., Kerwin D.R., Williamson J.D., Marottoli R.A., Rushing S., Marsiske M., Rapp S.R., LIFE Study Group: Performance of a computer-based assessment of cognitive function measures in two cohorts of seniors *Int. J. Geriatr. Psychiatry*, 28, pp. 1239–1250 (2013)
- [32] Brucki S.M.D., Nitrini R., Caramelli P., Bertolucci P.H.F., Okamoto I.H.: [Suggestions for utilization of the mini-mental state examination in Brazil] *Arq. Neuropsiquiatr.*, 61, pp. 777–781 (2003)
- [33] Nitrini R., Lefèvre B.H., Mathias S.C., Caramelli P., Carrilho P.E., Sauer N., Massad E., Takiguti C., Da Silva I.O., Porto C.S.: [Neuropsychological tests of simple application for diagnosing dementia] *Arq. Neuropsiquiatr.*, 52, pp. 457–465 (1994)
- [34] Dickson A.L.: Book Review: Lezak, Muriel Deutsch, (1983). *Neuropsychological assessment* (2nd ed.). New York: Oxford University Press J. Psychoeduc. Assess., 4, pp. 91–92 (1986)
- [35] Kaplan E., Goodglass H., Weintraub S.: Boston Naming Test, <http://dx.doi.org/10.1037/t27208-000>
- [36] Fuzikawa C., Lima-Costa M.F., Uchoa E., Barreto S.M., Shulman K., Bambuí Health and Ageing Study: A population-based study on the intra and inter-rater reliability of the clock drawing test in Brazil: the Bambuí Health and Ageing Study *Int. J. Geriatr. Psychiatry*, 18, pp. 450–456 (2003)
- [37] Brucki S.M.D., Rocha M.S.G.: Category fluency test: effects of age, gender and education on total scores, clustering and switching in Brazilian Portuguese-speaking subjects *Braz. J. Med. Biol. Res.*, 37, pp. 1771–1777 (2004)
- [38] Mols J.C., Heyman A., Mohs R.C., Hughes J.P., van Belle G., Fillenbaum G., Mellits E.D., Clark C.: The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease *Neurology*, 39, pp. 1159–1159 (1989)
- [39] Roth M., Huppert F.A., Tym E., Mountjoy C.Q., Diffident-Brown A., Shoesmith D.J.: The Cambridge Examination for Mental Disorders of the Elderly: CAMDEX, Cambridge University Press, (1988)
- [40] DSM-5, <https://www.psychiatry.org/psychiatrists/practice/dsm>
- [41] Atkinson M.T., Gucukoglu S., Machin C.H.C., Lawrence A.E.: Making the Mainstream Accessible: What's in a Game? *Lecture Notes in Computer Science*, pp. 380–387 (2006)
- [42] Grammenos D.: Game over Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08 (2008)
- [43] Bierre K., Chetwynd J., Ellis B., Michelle D., Ludi S., Westin T.: Game Not Over: Accessibility Issues in Video Games *Proc. of the 3rd International Conference on Universal Access in Human-Computer Interaction* (2005)
- [44] How WAI Develops Accessibility Guidelines through the W3C Process: Milestones and Opportunities to Contribute. Web Accessibility Initiative. W3C, <https://www.w3.org/WAI/intro/w3c-process.php>
- [45] Yuan B., Folmer E., Harris F.C.: Game accessibility: a survey *Universal Access in the Information Society*, 10, pp. 81–100 (2010)

Authoring of interventions with on-demand screen generation for mobile devices

Bruna C. R. Cunha^{1,2}, Kamila R. H. Rodrigues¹, Leonardo F. Scalco¹, Isabela Zaine¹, Caio C. Viel¹,
Maria da Graça C. Pimentel¹

Universidade de São Paulo¹, São Carlos, Brazil

Instituto Federal de São Paulo², São Paulo, Brasil

{brunaru7,kamila.rios,leoscalcoxe,isabela.zaine,caioviel}@gmail.com,mgp@usp.br

ABSTRACT

In this paper we present the ESPIM system, a set of tools that enables the authoring of intervention programs with focus on health and education. Created programs are played in a mobile app that generates screens on demand, according to their definition provided by JSON files. Using state-of-art technology we aim to provide a platform that can be used by professionals of health and education with no programming skills. To achieve that the system presents an intuitive authoring interface, built with requirements obtained with experts in different areas and with continuous feedback from evaluations, and a friendly mobile app, which presents elements that users are familiar and a good aesthetic design to motivate them. The system is being successfully used by professionals in health-related areas and shows many advantages in comparison to similar platforms.

KEYWORDS

ESM, mobile interventions, authoring, automatic screen generation

1 INTRODUÇÃO

A crescente democratização do uso de *smartphones* tornou a área de *mhealth*, a qual utiliza dispositivos móveis para ofertar serviços de saúde, uma área de pesquisa fortemente presente na realidade das pessoas, promovendo a diminuição da distância entre pesquisadores, profissionais de saúde e população. O sucesso e popularização da área é constatado pela grande disponibilidade de aplicativos relacionados nas lojas dos sistemas Android e iOS, assim como a grande quantidade de artigos científicos [12] publicados nos últimos anos.

Pesquisas envolvendo aplicativos de saúde para dispositivos móveis demonstram que o uso de intervenções pode aumentar a motivação de indivíduos [10], além disso, o uso de dispositivos móveis para promover intervenções é capaz de diminuir disparidades de saúde entre grupos mais vulneráveis [13].

No contexto da saúde, a coleta de informações e o provimento de informações personalizadas para indivíduos com demandas específicas permite aprimorar o tratamento de pacientes. Porém, a utilização de papel para este fim é muitas vezes pouco eficiente, dada a dificuldade de gerenciamento de informação e baixa aderência de pacientes. A utilização de dispositivos móveis em diferentes

cenários de tratamento já demonstrou ser eficiente em aumentar a aderência de participantes [10].

Intervenções personalizadas também apresentam utilidade em outros contextos, como o educacional. A utilização de dispositivos móveis no ensino à distância, designado pelo termo *m-learning*, e a criação de conteúdo sob demanda para diferentes perfis de alunos, tem se mostrado uma estratégia eficiente no aprendizado, promovendo maior motivação e engajamento [16].

O Método de Amostragem de Experiências (ESM) [6] é um método verificado e bem-sucedido para a coleta dados sobre as pessoas em seu ambiente natural e, apesar de proposto em papel, seu uso em dispositivos computacionais tornou-se uma alternativa viável. Já no contexto interventivo, a abordagem de Sistema Personalizado de Instrução (PSI) [8] mostrou-se eficiente em melhorar o aprendizado. Com o objetivo de criar um método que permita coletar informações e gerar conteúdo instrucional personalizado via dispositivos móveis, em trabalhos anteriores, os autores propuseram o Método de Amostragem de Experiências e Intervenção Programada (ESPIM) [14] que faz uso da Computação Ubíqua apoiada pelos métodos ESM e PSI. As ferramentas apresentadas neste trabalho são apoiadas pelo método ESPIM.

O conjunto de ferramentas busca prover a autoria de programas interventivos, que podem ser compostos por intervenções de coleta ou de conteúdos instrucionais, por parte de profissionais e especialistas de saúde e educação, assim como a reprodução de telas interativas em dispositivos móveis para a interação com os públicos alvo. Assim foram desenvolvidos um sistema Web para a autoria dos programas de intervenção personalizada que gera documentos JSON que representam os programas e um aplicativo para *smartphones* e *tablets* com sistema Android, que recebe os documentos e gera dinamicamente as telas de interação, bem como as notificações disparadas em horários definidos durante a autoria. Este trabalho objetiva detalhar as ferramentas citadas.

2 TRABALHOS RELACIONADOS

No contexto de *mhealth*, muitos aplicativos interventivos foram propostos com o intuito de coletar dados [18], disparar lembretes de medicação [5], aumentar a motivação e aderência em atividades [4], enviar instruções ou informações [9], ou uma combinação destas funcionalidade para realizar acompanhamentos em condições específicas [1, 11]. No entanto, esses aplicativos foram implementados por profissionais de computação, seguindo requisitos de domínios específicos, sem permitir a flexibilidade de customização por profissionais de outros domínios.

Considerando aplicações computacionais que utilizam a metodologia ESM para coletar dados, Berkel *et al.* [3] apontam a falta

In: XVII Workshop de Ferramentas e Aplicações (WEA 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pôsteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.

© 2018 SBC – Sociedade Brasileira de Computação.

ISBN 978-85-7669-435-9.

de ferramentas que permitam a autoria de experimentos por parte de especialistas. A criação de experimentos que utilizam ESM e dispositivos móveis está bastante atrelada a necessidade de um profissional com conhecimentos de programação.

Procurando apresentar uma forma de personalização, o *framework* AWARE [7] permite a criação de questões ESM personalizadas, no entanto, sua programação deve ser feita em um arquivo JSON. A ExperienceSampler [17], apesar de afirmar propor uma interface de autoria simplificada, exige que o profissional execute uma série linhas de comandos e também faça suas configurações em um arquivo de texto. A plataforma PACO [2] permite que pesquisadores realizem a autoria de experimentos personalizados por meio de um Website e que usuários se registrem em experimentos em seus dispositivos móveis. Sua interface de autoria contém termos da área da Computação, principalmente relacionada à programação, e configurações complexas, o que dificulta seu entendimento para leigos. A ferramenta Ohmage [15], por sua vez, permite criar experimentos personalizados com questões específicas, porém apresenta pouca flexibilidade em termos de configuração de intervenções, não permitindo, por exemplo, a visualização e captura de mídias.

A *ilumivu*¹ é uma plataforma paga que, apesar de apresentar um aplicativo móvel com interface convidativa, possui uma interface de autoria complexa e limitada, além de apenas permitir questões com respostas em escala ou texto. O *movisensXS*² é um sistema que permite a criação de experimentos ESM que pode ser considerado um sistema que representa bem o estado-da-arte, no entanto sistema é pago e não possui código aberto, de forma que é pouco customizável ou extensível.

As ferramentas citadas apresentam limitações em termos de interação com os usuários, tanto na autoria quanto na reprodução. Este trabalho visa apresentar uma alternativa mais completa no contexto de criação de programas interventivos, que permita a criação de conteúdos textuais e multimídia, assim como a solicitação de resposta em formato multimídia.

3 ESPIM: SISTEMAS PARA AUTORIA E REPRODUÇÃO DE INTERVENÇÕES

O sistema ESPIM utiliza o Design Participativo como estratégia de desenvolvimento, de modo que especialistas das áreas da computação, saúde e educação estão envolvidos desde a etapa de levantamento de requisitos até as fases de avaliação e validação. O sistema é composto por um modelo e por duas ferramentas, sendo uma de autoria e outra de reprodução. É importante destacar que as interfaces das ferramentas aqui apresentadas contaram com o apoio de designers experientes na criação de interfaces gráficas, de forma que, além de aprimorar a experiência do usuário, o design atual apresenta padrões estéticos de um produto comercial. As próximas seções buscam detalhar a arquitetura do sistema e suas ferramentas.

3.1 Aplicação Web para Autoria

O objetivo da ferramenta de autoria é permitir que profissionais, especialistas das áreas de saúde e educação, elaborem programas de intervenção personalizados e interativos, que contenham texto, questões e mídias, para serem exibidos nos dispositivos móveis de

participantes cadastrados. Além de criar os programas, a ferramenta permite que os profissionais visualizem os resultados das interações de seus participantes com os programas em seus dispositivos.

A ferramenta de autoria trata-se de uma *Rich Internet Application* (RIA), desenvolvida fazendo o extenso uso de JavaScript. A ferramenta de autoria também contém um módulo de visualização de resultados. Toda a aplicação foi desenvolvida utilizando HTML5, CSS e o *framework* CSS Bootstrap, seguindo princípios de *Material Design*³ e design responsivo. A interface de autoria utiliza o *framework* JavaScript Angular.js, enquanto a de visualização de resultados utiliza os *frameworks* JavaScript React, D3.js e plotly.js, sendo os dois últimos voltados especificamente para a visualização de dados e geração de gráficos respectivamente.

A interface de autoria permite aos especialistas criar, exportar e importar programas personalizados para coleta de dados, para disponibilização de informações ou para intervenção remota de participantes, os quais devem ser incluídos nos programas de interesse. Esses programas também podem ter observadores, tais como outros colegas profissionais, parentes ou cuidadores. Um programa é composto por um conjunto de intervenções, agrupadas em uma entidade denominada evento. Um evento contém disparos temporais que podem ser de diferentes tipos: uma notificação discreta, uma notificação longa ou um alarme. Um participante também pode iniciar a reprodução manualmente clicando no evento listado no aplicativo. Durante a configuração do evento o especialista pode bloquear ou permitir a inicialização manual pelo participante.

As intervenções são as unidades de interação com o usuário final, ou seja, representam o conteúdo exibido e requisitado aos participantes em seus dispositivos móveis. Os seguintes tipos de intervenção estão disponíveis para criação: questões abertas, questões de múltipla escolha ou escolha única, mensagens, solicitação de mídias (áudio, vídeo, imagem) e aplicação externa (em que um aplicativo externo pode ser iniciado). O enunciado das intervenções pode conter mídias do tipo imagem, áudio e vídeo, enviadas pelo especialista via interface de gravação e *upload* de mídias.

A interface de autoria de intervenções utiliza uma tela semelhante a um fluxograma dinâmico, que oferece uma maneira flexível e visual de conectá-las, criando fluxos sequenciais. A Figura 1 ilustra a tela para a criação das intervenções usando a interface Web de autoria do sistema.

A aplicação Web dispõe de um módulo para a visualização e análise dos dados coletados, o qual pode ser acessado por qualquer observador do programa de interesse. Na área de resultados, após selecionar um programa de interesse, o observador recebe uma tela que apresenta a quantidade de resultados recuperados até o momento e o número de execuções do evento por participantes. Nesta tela o observador tem a opção de baixar um arquivo do tipo Comma Separated Value (CSV), que contém todos os dados coletados referentes às respostas de todos os participantes, ou realizar uma visualização detalhada dos dados por participante. Essa última tela contém recursos gráficos para análise de resultados.

A visualização por participante disponibiliza o detalhamento de suas interações com as intervenções. Com base nos objetos retornados pelo documento JSON, oriundos das respostas aos programas, quatro representações gráficas são exibidas. Elas disponibilizam

¹<https://ilumivu.com/>

²<https://xs.movisens.com/>

³<https://material.io/design/>

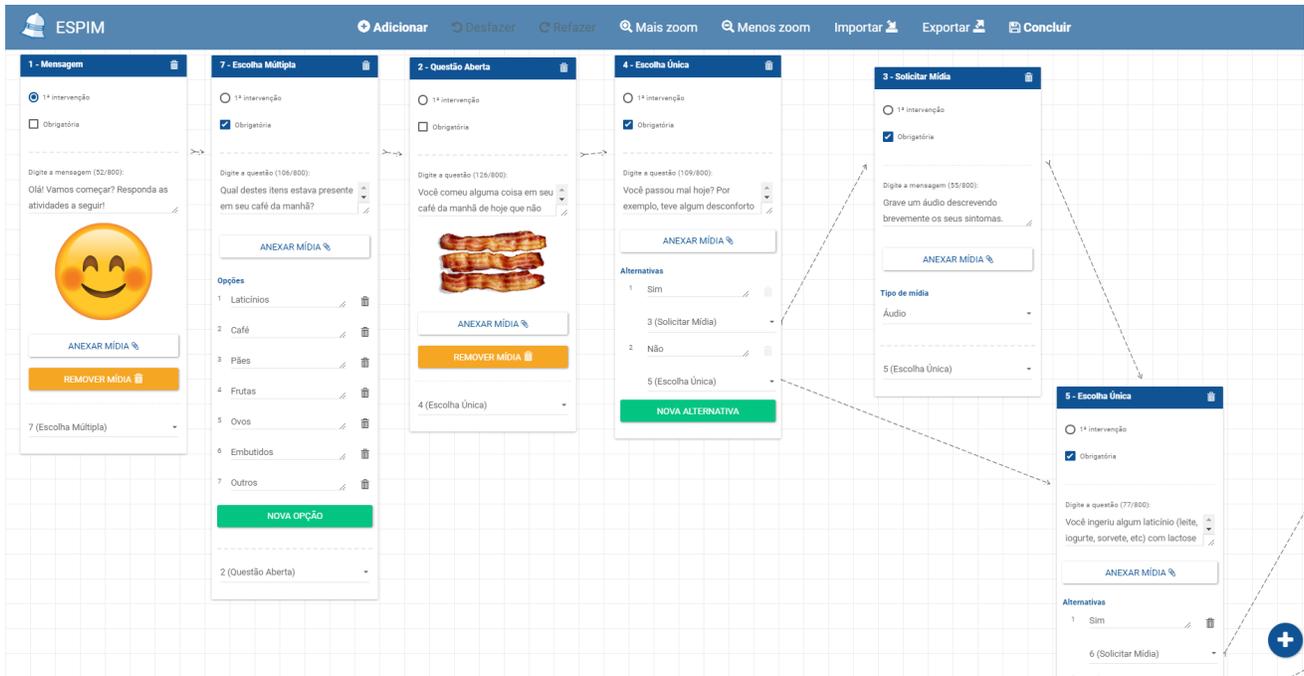


Figura 1: Exemplo da interface de autoria de intervenções.

informações tais como: em qual dia determinado participante mais respondeu às intervenções planejadas, em qual período do dia essa interação aconteceu, quais são as intervenções mais respondidas e quais intervenções os participantes levam mais tempo para concluir a interação. Além dos gráficos, a tela possui uma tabela com todos os resultados coletados. Na tabela, para cada intervenção é exibido seu tipo, horário do início da interação, horário do fim da interação, o enunciado da intervenção e o resultado enviado pelo usuário, caso aplicável.

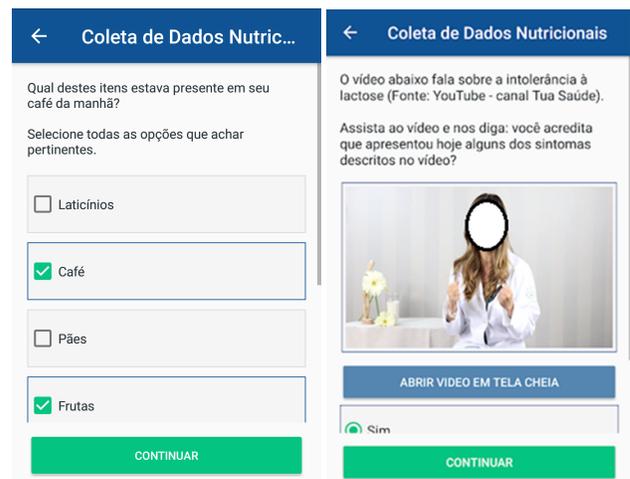
A aplicação para autoria de programas ESPIM está disponível em: <https://goo.gl/YPvwmb>.

3.2 Aplicativo Móvel para Reprodução

O aplicativo Smart ESPIM aNd Sensor Mobile (SENSEM) é responsável por recuperar os programas ESPIM criados via aplicação Web e gerar sob demanda as telas de intervenções, de acordo com o que foi configurado durante a autoria. Em sua primeira execução o aplicativo apresenta telas de apresentação e pede para que o usuário realize o login usando a sua conta Google. Após o login, é realizada uma requisição REST para verificar se o usuário está cadastrado em algum programa e, em caso positivo, os programas são recuperados. O Web Service retorna ao aplicativo um documento JSON que contém todos os programas em que esse usuário foi cadastrado e é com base neste documento que as telas são geradas.

Como já citado, as telas são baseadas nas seguintes possibilidades de intervenções: mensagens (as quais apresentam apenas conteúdo), questões abertas, questões de múltipla escolha com única ou múltipla seleção, captura de mídias (foto, áudio e vídeo) e inicialização de aplicações externas. Todas as telas seguem um determinado *template* estruturado, de forma que o texto aparece

sempre no topo da tela e em seguida é apresentada a mídia opcional, caso presente. Abaixo deste conteúdo é apresentada a interação que o usuário deve realizar, que pode ser obrigatória ou não. A Figura 2 ilustra exemplos de telas do aplicativo.



(a) Enunciado textual

(b) Enunciado multimídia

Figura 2: Telas do aplicativo com questões de múltipla escolha e diferentes enunciados.

Intervenções que solicitam a captura de mídias são apresentadas de forma diferenciada, de acordo com a mídia solicitada para facilitar a interação com usuário. No caso de vídeo e foto é exibida uma tela

similar que invoca a câmera para gravação (Figura 3 (a)). Já para a gravação de áudio é exibido um botão com a instrução “Segure para gravar” (Figura 3 (b)). Após a primeira captura, a mídia capturada pode ser reproduzida na própria tela e o usuário pode optar por capturar novamente caso não fique satisfeito com o resultado.

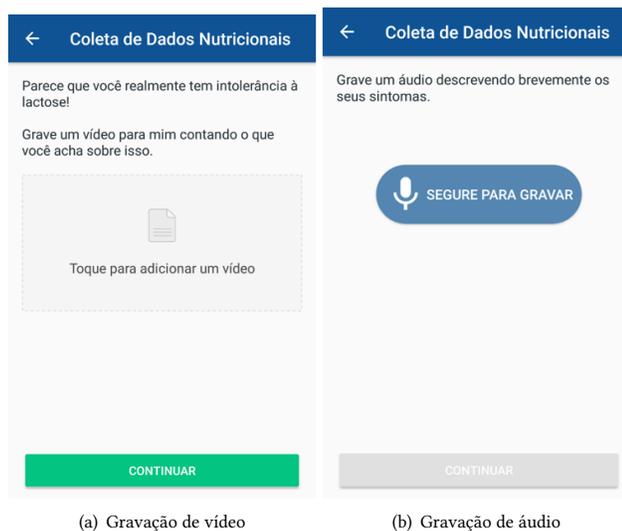


Figura 3: Telas do aplicativo para captura de mídias.

Durante a interação é verificada a obrigatoriedade das intervenções, assim como as opções que podem levar a diferentes fluxos. O usuário pode voltar e alterar o valor de respostas anteriores, o que pode alterar o fluxo posterior. Para isso, foi implementado um algoritmo que realiza o controle do caminho percorrido e das respostas válidas respondidas durante a interação.

Os disparos temporais, configurados no evento pela interface de autoria, apresentam três formatos: uma notificação com som padrão do sistema, uma notificação com um alerta sonoro de dez segundos e um alarme, o qual abre a aplicação na tela e toca um alerta sonoro por um minuto ou até o usuário interagir com o aplicativo.

Quando o usuário finaliza a interação com um conjunto de intervenções, os dados da interação são armazenados localmente no dispositivo e, caso haja conexão com a Internet, enviados instantaneamente para o Web Service, caso contrário, o aplicativo enviará os dados quando houver uma conexão disponível. O aplicativo também apresenta funcionalidades de limpeza automática e manual de mídias, de forma a não sobrecarregar o armazenamento dos dispositivos.

O aplicativo SENSEM para reprodução de programas ESPIM está disponível em: <https://goo.gl/yDXUrp>. O SENSEM é compatível com *smartphones* e *tablets* com sistema operacional Android à partir da versão KitKat (4.4) e é compilado e testado na sua versão mais recente (8.0, Oreo)⁴.

⁴<https://www.android.com/versions/oreo-8-0/>

3.3 Arquitetura

A arquitetura do sistema buscou criar independência entre seus componentes e dados. Desta forma, a arquitetura REST e o uso de armazenamento de dados na nuvem foram as estratégias escolhidas. Assim, toda a comunicação entre os componentes é realizada por meio da Internet, como mostra a Figura 4.

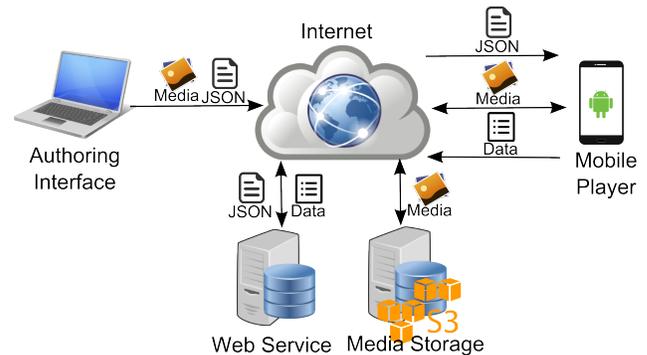


Figura 4: Arquitetura simplificada do sistema.

O modelo de dados para o método ESPIM foi concebido considerando um conjunto amplo de requisitos que foram elicitados, inicialmente, por uma revisão da literatura de artigos que abordavam ESM e ferramentas *mhealth* e, mais extensivamente, por entrevistas com profissionais da área da saúde e educação, as quais geraram uma grande variedade de particularidades, estas previstas no modelo de forma a permitir generalização e extensão das possibilidades de construção dos programas.

Os dados de programas interventivos, armazenados no formato do modelo proposto, são acessados por um Web Service REST, de forma que a comunicação entre as ferramentas é feita pela troca de documentos JSON. As mídias no formato de áudio, imagem e vídeo são armazenadas em um servidor independente. Neste caso, é utilizado o serviço Amazon S3⁵, devido sua alta capacidade de armazenamento. Ao salvar um programa na ferramenta de autoria, o documento JSON resultante é enviado ao Web Service, enquanto as mídias são enviadas para o servidor de mídias. O documento JSON possui apenas os endereços URL das mídias armazenadas, de forma que o documento transporta apenas informação textual.

O aplicativo Android realiza requisições para recuperar documentos JSON referentes aos programas em que o usuário está cadastrado. As mídias são recuperadas por meio de suas URLs disponíveis nos documentos. O aplicativo também envia documentos JSON, que contém os resultados das interações realizadas pelos usuários, para serem armazenados pelo Web Service, e mídias capturadas para o servidor S3.

Atualmente o cadastro e login no sistema, tanto para autores quanto para reprodutores de programas, é realizado por meio da conta Google deste usuário. No caso do especialista, autor de programas, o cadastro é feito na primeira vez que acessa o sistema, utilizando a opção de *Sign in with Google* na primeira página da aplicação Web. Os especialistas precisam dos endereços de e-mail Google dos participantes para cadastrá-los nos programas. Já no

⁵<https://aws.amazon.com/s3/>

aplicativo, dada a integração da conta Google com o sistema Android, o login é realizado ao selecionar a conta desejada. Esta abordagem foi utilizada por questões de segurança, já que utiliza um sistema seguro para gerenciamento de login e senha.

4 AVALIAÇÕES

Durante o processo de concepção do sistema ESPIM foram realizadas avaliações em suas ferramentas com a colaboração de especialistas de diferentes áreas. Entre as avaliações realizadas estão: testes empíricos com usuários, inspeções heurísticas de usabilidade e um teste de usabilidade com usuários. Ressalta-se aqui, que este projeto foi aprovado pelo Comitê de Ética em Pesquisa com número de processo 57875016.3.0000.5390.

A interface Web de autoria está em sua 5ª versão e foi submetida a uma avaliação heurística com pesquisadores da área Interação Humano-Computador e a um teste de usabilidade com usuários profissionais das áreas de terapia ocupacional e gerontologia. No momento o time concentra esforços em oferecer recursos de ajuda e documentação para a ferramenta, tais como vídeos instrucionais e uma Wiki.

O aplicativo SENSEM está em sua 3ª versão e passou por duas avaliações de inspeção heurística, além de estar sendo utilizado por diferentes especialistas (gerontólogos, terapeutas ocupacionais e psicólogos) em intervenções e coletas de dados com suas populações de interesse. Entre as coletas cita-se: lembretes de atividades para pais de crianças autistas (psicologia), reforço em cursos de letramento digital para idosos (terapia ocupacional e computação), autoria de instruções para cuidadores (gerontologia) e compartilhamento de mídias para aprimorar relacionamento entre pessoas idosas (psicologia). A condução destes experimentos com usuários reais foi de extrema importância para revelar *bugs* de *software* e problemas de usabilidade, os quais foram corrigidos no decorrer de sua execução. No momento o time está trabalhando na disponibilização de disparos contextuais de eventos.

5 CONCLUSÃO

A versão atual do sistema representa um esforço conjunto do time de pesquisadores e desenvolvedores que compõe a equipe deste projeto, na busca constante pelo aperfeiçoamento de toda a infraestrutura. Os desenvolvedores têm realizado manutenções corretivas e evolutivas no sistema a partir de testes e do *feedback* coletado com as avaliações empíricas e analíticas realizadas.

O sistema ESPIM tem sido usado por terapeutas ocupacionais, gerontólogos, psicólogos e professores para a coleta de dados e intervenções com suas populações de interesse, tanto em suas práticas profissionais, quanto de pesquisa. A adesão e uso do sistema por esses e outros profissionais revelam o potencial do mesmo para atividades diversas, incluindo atividades de monitoramento, compartilhamento de mídias, informes à população, entre outros.

O objetivo do grupo de pesquisa é fornecer o sistema no formato de *software* livre, para que outros grupos de pesquisa e profissionais possam usar e/ou modificar o sistema para atender aos seus interesses profissionais ou de pesquisa. Os códigos das ferramentas serão distribuídos sob a licença GNU *General Public License* (GPLv3)⁶. Um

tutorial mostrando a criação e exibição de um programa ESPIM está disponível em <https://youtu.be/wcly81tGjQ>.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da CAPES, FAPESP e CNPq.

REFERÊNCIAS

- [1] Ana Correia Barros, João Cevada de, Âgels Bayés, Sheila Alcaine, and Berta Mestre. 2013. User-centred Design of a Mobile Self-management Solution for Parkinson's Disease. In *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia*. ACM, Article 23, 10 pages.
- [2] Kathy K Baxter, Anna Avrek, and Bob Evans. 2015. Using experience sampling methodology to collect deep data about your users. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 2489–2490.
- [3] Niels Van Berkel, Denzil Ferreira, and Vassilis Kostakos. 2017. The Experience Sampling Method on Mobile Devices. *Comput. Surveys* 50, 6 (2017), 93.
- [4] Paulina Bondaronek, Elizabeth Murray, and Fiona Hamilton. 2017. The Public Health Potential of the Current Health Apps for Increasing Physical Activity. In *Proceedings of the 2017 International Conference on Digital Health (DH '17)*. ACM, New York, NY, USA, 229–230.
- [5] Nichapat Boonnudat and Pongpisit Wuttidittachotti. 2017. Mobile Application: Patients' Adherence to Medicine In-take Schedules. In *Proceedings of the International Conference on Big Data and Internet of Thing (BDIoT2017)*. ACM, New York, NY, USA, 237–241.
- [6] M. Csikszentmihalyi and R. Larson. 1987. Validity and reliability of the Experience-Sampling Method. *The Journal of Nervous and Mental Disease* 175, 9 (Sept. 1987), 526–536.
- [7] Denzil Ferreira, Vassilis Kostakos, and Anind K Dey. 2015. AWARE: mobile context instrumentation framework. *Frontiers in ICT* 2 (2015), 6.
- [8] Fred S. Keller. 1968. "Good-bye, teacher...". *Journal of Applied Behavior Analysis* 1, 1 (1968), 79–89.
- [9] Pei-Yi (Patricia) Kuo and Michael S. Horn. 2017. Daily Challenges for Sustainable Lifestyles: Design Implications from a Mobile Intervention Study. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers (UbiComp '17)*. ACM, New York, NY, USA, 635–641.
- [10] Nicholas D. Lane, Mu Lin, Mashfiqui Mohammad, Xiaochao Yang, Hong Lu, Giuseppe Cardone, Shahid Ali, Afsaneh Doryab, Ethan Berke, Andrew T. Campbell, and Tanzeem Choudhury. 2014. BeWell: Sensing Sleep, Physical Activities and Social Interactions to Promote Wellbeing. *Mobile Networks and Applications* 19, 3 (June 2014), 345–359.
- [11] David C. Mohr, Enid Montague, Colleen Stiles-Shields, Susan M. Kaiser, Christopher Brenner, Eric Carty-Fickes, Hannah Palac, and Jenna Duffecy. 2015. MedLink: A Mobile Intervention to Address Failure Points in the Treatment of Depression in General Medicine. In *Proceedings of the 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth '15)*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), ICST, Brussels, Belgium, Belgium, 100–107.
- [12] Editorial Office. 2017. mHealth is now indexed in PubMed/PMC: a tremendous step forward. *mHealth* 3 (March 2017).
- [13] Monica E. Peek. 2017. Can mHealth Interventions Reduce Health Disparities among Vulnerable Populations? *Diversity & Equality in Health and Care* 14, 2 (March 2017).
- [14] Kamila R.H. Rodrigues, Caio C. Viel, Isabela Zaine, Bruna C.R. Cunha, Leonardo F. Scalco, and Maria G.C. Pimentel. 2017. Data Collection and Intervention Personalized As Interactive Multimedia Documents. In *Proceedings of the 23rd Brazilian Symposium on Multimedia and the Web*. ACM, New York, NY, USA, 57–60.
- [15] Hongsuda Tangmunarunkit, Cheng-Kang Hsieh, Brent Longstaff, S Nolen, John Jenkins, Cameron Ketcham, Joshua Selsky, Faisal Alquaddoomi, Dony George, Jinha Kang, et al. 2015. Ohmage: A general and extensible end-to-end participatory sensing platform. *ACM Transactions on Intelligent Systems and Technology* 6, 3 (2015), 38.
- [16] Liane MR Tarouco, MCJM Fabre, Anita RS Grando, and Mary LP Konrath. 2004. Objetos de Aprendizagem para M-learning. In *Florianópolis: SUCEU-Congresso Nacional de Tecnologia da Informação e Comunicação*.
- [17] Sabrina Thai and Elizabeth Page-Gould. 2017. ExperienceSampler: An Open-Source Scaffold for Building Smartphone Apps for Experience Sampling. (2017).
- [18] Zhi Zheng, Loren Bruns, Jr, Jiajie Li, and Richard O. Sinnott. 2017. A Mobile Application and Cloud Platform Supporting Research into Alcohol Consumption. In *Proceedings of the 1st International Conference on Medical and Health Informatics 2017 (ICMHI '17)*. ACM, New York, NY, USA, 48–55.

⁶<https://www.gnu.org/licenses/gpl-3.0.html>

Ballgorithm - Uma Ferramenta Introdutória Para Conceitos de Programação

Guilherme Afonso Melo Sousa Melo
Laboratório Telemídia - MA
Av. dos Portugueses, 1966
São Luís, MA 65080-805
guilhermemelo6@outlook.com

Carlos de Salles Soares Neto
Laboratório Telemídia - MA
Av. dos Portugueses, 1966
São Luís, MA 65080-805
csalles@deinf.ufma.br

ABSTRACT

The Computer Science Course presents one of the highest dropouts rates among higher education courses in Brazil. Besides that, it's known that the Algorithms discipline is the base of the knowledge learnt throughout the course and the first contact of students with computer programming. Having that said, the Ballgorithm was idealized to be a tool that includes a simple programming language and a ludical approach, making use of digital games visual elements. This paper aims to present the development process of the Ballgorithm tool, the Ballcode language and the design decisions regarding to the system architecture.

KEYWORDS

Serious Games. Gamification. Teaching.

1 INTRODUÇÃO

Um problema recorrente entre os cursos de graduação em Ciência da Computação é a alta taxa de desistência. A taxa de estudantes que concluem o curso é de 14.3% [5]. Além de ser um número baixo por si só, torna-se ainda mais preocupante se comparado com a média geral de conclusões em universidades públicas, que é 45.9% [4]. Segundo Santos e Costa [1], o ensino de algoritmos acaba sendo a base do conhecimento durante o curso, e a dificuldade de aprendizado deste conteúdo acaba tornando-se um dos principais motivos para desistência do curso.

Dessa forma, é fundamental que o ensino de algoritmos seja tratado com relevante grau de importância, já que, além de ser a base dos conhecimentos que serão aprendidos ao longo da graduação, acaba sendo também o primeiro contato de muitos dos ingressantes com programação de computadores. Tendo isto em vista, é importante buscar métodos alternativos de ensino que aumentem a motivação do aluno, os quais sejam mais dinâmicos e lúdicos [3].

O objetivo deste trabalho é apresentar a ferramenta *Ballgorithm*, idealizada com o intuito de trazer uma abordagem mais lúdica ao aprendizado de algoritmos, apresentando atividades simples e dinâmicas, bem como a representação visual dos algoritmos como um jogo. Na Seção 2 apresentaremos a estrutura da *Ballgorithm*, falando de seus componentes e de como a ferramenta pode ser usada pelos usuários. Na Seção 3 entramos em detalhes sobre o processo de interpretação da linguagem e exibição dos objetos em tela. Por

In: XVII Workshop de Ferramentas e Aplicações (WFA 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pôsteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.
© 2018 SBC – Sociedade Brasileira de Computação.
ISBN 978-85-7669-435-9.

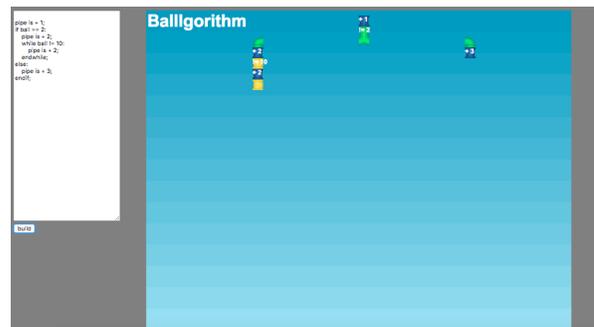
fim, na Seção 4 falamos brevemente sobre resultados preliminares obtidos através de testes com usuários.

2 BALLGORITHM

O *Ballgorithm* é uma ferramenta que tem como objetivo a prática e o aprendizado de conceitos básicos de linguagens de programação. Por conta disso, a ferramenta (que pode ser qualificada como um jogo sério) conta com sua própria linguagem, chamada *Ballcode*. O jogo funciona da seguinte forma: desafios envolvendo os elementos em tela (canos com operações matemáticas e bolas compostas por valores numéricos) devem ser resolvidos pelo jogador, criando um percurso com canos que irão alterar os valores numéricos presentes nas bolas. Ao final deste percurso, as bolas devem ter os seus valores equivalentes ao valor solicitado.

Na Figura 1, podemos ver a tela do jogo, que foi gerada após a submissão do código escrito pelo usuário.

Figure 1: Tela do Ballgorithm



O jogo consiste em se levar uma ou mais bolas do topo da tela a um ponto na base da tela, fazendo com que, ao final do percurso, os valores numéricos das bolas sejam iguais ao valor solicitado. Os valores especificados podem ser declarados tanto dentro dos níveis do jogo (pré-programados por outros usuários) quanto propostos por outra pessoa (professores, amigos, etc).

Para alterar estes valores, o jogador precisa utilizar os diferentes tipos de canos que se encontram disponíveis através da linguagem *Ballcode*, criada especificamente para o jogo. Cada cano tem uma função diferente, e parte do desafio presente no jogo é entender como alocá-los de forma eficiente, levando-se em consideração o número de canos utilizados. É importante ressaltar também o potencial de competitividade entre os usuários, tendo em vista que

a pontuação gerada ao término de um nível leva em consideração o número de canos utilizados na solução.

Assim sendo, pode-se afirmar que o *Ballgorithm* é, na verdade, uma ferramenta que simula uma máquina formada por canos, capaz de efetuar o processamento em sua entrada (bolas com valores numéricos quaisquer) e retornar de forma bem sucedida uma saída (bolas com os valores desejados).

Buscando uma forma mais simples de demonstrar na prática os conceitos básicos ensinados em programação, chegou-se no modelo atual do jogo. Note que os conceitos básicos de programação devem ser apresentados a usuários que tenham pouco (ou nenhum) conhecimento prévio. Dessa forma, os seguintes conceitos foram escolhidos:

- Entrada e Saída
- Atribuições
- Estruturas Condicionais
- Estruturas de Repetição

Estes conceitos podem ser percebidos nas abstrações presentes nos objetos que podem ser utilizados na ferramenta: a criação de novas bolas representa a entrada; os canos de atribuições; estruturas condicionais e de repetição fazem as vias de unidade de processamento; enquanto o valor final da bola nada mais é do que a saída da máquina criada pelo usuário utilizando tais elementos.

2.1 A Linguagem *Ballcode*

A linguagem *Ballcode* foi criada especificamente para o *Ballgorithm*, e o ponto mais importante da sua criação foi a sua simplicidade, já que é importante que toda (ou a maior parte) da dificuldade relacionada à resolução dos desafios fosse proveniente exclusivamente do raciocínio lógico envolvido nas questões. Dessa forma, optou-se por criar uma linguagem com sintaxe baseada em Python, por essa ser uma linguagem de fácil leitura e pelo foco que a mesma tem na indentação do código escrito.

2.1.1 Comandos da Linguagem.

- (1) **Declarar Instância de uma nova bola:** O comando de atribuição atribui um valor a uma bola que será criada. Tal comando só atribui valores numéricos a bolas que estão sendo instanciadas. Ou seja, não é possível alterar o valor de bolas já existentes no jogo através do comando de atribuição.
- (2) **If/ Estruturas Condicionais:** O comando de estruturas condicionais tem como objetivo definir uma bifurcação no percurso que será feito pelas bolas. As bolas entram no cano da estrutura condicional e, baseando-se no seu valor numérico atual, define-se qual caminho será percorrido. Diferentemente das estruturas semelhantes nas mais diversas linguagens de programação, que checam se os valores são iguais, diferentes, maior que e menor que, as estruturas de controle em *Ballcode* checam apenas se os valores comparados são iguais ou diferentes. Na Figura 2, vemos duas estruturas condicionais aninhadas, onde, no primeiro caso, se o valor da bola for diferente de 5 (cinco), esta segue pra direita, e caso contrário, ela segue para a esquerda. Dentro deste cano existem duas operações, onde o valor da bola é multiplicado por 4 (quatro) e depois dividido por 4

(quatro). Depois disso, há mais uma estrutura condicional, onde, caso o cano seja diferente de 1 (um), ele seguirá para a direita, e caso contrário, para a esquerda.

Figure 2: Bloco de canos gerados pelo comando das Estruturas Condicionais



- (3) **Atribuições de Valores/Operações Básicas:** As atribuições de valores permitem que o usuário altere o valor das bolas que foram declaradas anteriormente, através de operações matemáticas básicas, como adição, subtração, multiplicação e divisão. A atribuição de valores é a principal ferramenta para a resolução dos problemas propostos. Na figura 3 podemos ver um exemplo de cano de atribuição.

Figure 3: Cano gerado pelo comando de atribuição



- (4) **While/Estrutura de repetição:** As estruturas de repetição permitem ao usuário utilizar um comando (ou blocos de comandos) repetidas vezes, de forma a evitar redundâncias e reduzir os caminhos de canos que serão montados. A estrutura de repetição no projeto conta com dois canos. Um cano determina a posição inicial do que seria o percurso a ser repetido e outro determina a posição final. Caso a condição especificada para a parada da repetição não seja cumprida, a bola, ao chegar na posição final do percurso, será mandada de volta para a posição inicial, até que a condição de parada seja atendida e, assim, ela possa continuar o trajeto.

Na Figura 4 podemos ver a estrutura que será gerada com o uso de estruturas de repetição, contendo uma atribuição dentro do bloco de código.

2.1.2 *Especificação da Sintaxe Ballcode.* Os comandos textuais utilizados no *Ballgorithm* são avaliados através de uma comparação com comandos pré-determinados existentes na linguagem. Estes comandos foram criados de forma a simular as operações mais básicas presentes na criação de algoritmos.

Abaixo, a Tabela 1 demonstra os comandos de *Ballcode*, bem como a sintaxe correta para utilizá-los:

Figure 4: Bloco de canos gerados pelo comando das Estruturas de repetição



Table 1: Comandos em Ballcode

Declaração de um novo objeto	new ball = [valor inteiro de um dígito];
Atribuição de um valor	pipe is [operação básica (+, -, * ou /)];
Estrutura condicional	if ball ['==' ou '!='] [valor inteiro de até 2 dígitos]: //bloco de código else: //bloco de código endif;
Estrutura de Repetição	while ball ['==' ou '!='] [valor inteiro de até 2 dígitos]: //bloco de código endwhile;

2.2 O Usuário enquanto agente ativo

Embora o *Ballgorithm* tenha sido idealizado inicialmente como uma ferramenta onde o usuário final não exerce nenhuma atividade além de resolver as questões solicitadas, viu-se que existe a possibilidade de o usuário submeter seus próprios desafios, tendo em vista que as máquinas podem ser modificadas de forma a manter determinado comando "trancado", de forma a fazer com que os usuários encontrem uma resposta para os enigmas sendo obrigados a jogar em condições mais limitadas, em vez de fazer com que a máquina seja gerada apenas com os comandos de sua escolha.

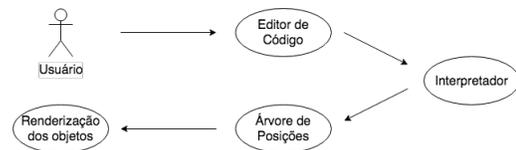
Como já foi descrito anteriormente, o *Ballgorithm* possui dois grandes pilares: a linguagem *Ballcode*, responsável por adicionar na tela os elementos utilizados no jogo, e o jogo propriamente dito, que exibe os objetos criados através da linguagem, além de contar com um sistema de física e movimentação para os objetos. Pode-se dizer que o funcionamento do *Ballgorithm* baseia-se em dois grandes componentes: o interpretador de *Ballcode* e o gerenciamento dos objetos presentes na tela do jogo. Além desses dois grandes componentes, também existem funções e objetos que têm como objetivo fazer a interação entre eles.

No momento em que o usuário decide enfim gerar o percurso descrito em seu código, o interpretador do *Ballgorithm* passa pelas seguintes etapas: leitura das linhas, normalização e classificação,

criação de objetos do tipo linha, definição das posições dos objetos na tela do jogo, criação de objetos presentes no jogo (*sprites*), instanciação dos *sprites* e, enfim, exibição do que foi gerado.

A Figura 5 ilustra, através de um diagrama de caso de uso, os processos que ocorrem entre a submissão do código do usuário e a renderização dos componentes do jogo em tela.

Figure 5: Diagrama de Caso de Uso



3 O BALLGORITHM E O COMPILADOR

Embora a linguagem *Ballcode* seja de fundamental importância para a experiência do usuário, é importante ressaltar que, de nenhuma forma, este projeto tem como objetivo funcionar como um compilador, embora muitas das funções existentes nele sejam análogas a certas funções presentes em um compilador. Considerando o modelo conceitual de um compilador, os passos por ele seguidos são análise do código, representação semântica e síntese, para então assim gerar o código executável [2].

De forma análoga ao compilador, o código do usuário, ao ser submetido, é analisado, gera uma lista com os objetos e sintetiza estes objetos naqueles que serão renderizados na tela do jogo, gerando assim o percurso descrito pelo usuário, o que seria, em uma comparação grosseira, o executável gerado pelo compilador.

3.1 O interpretador da linguagem *Ballcode*

O interpretador do *Ballgorithm* tem como objetivo extrair o conteúdo das linhas do código do usuário e, a partir destas, criar os objetos que podem ser organizados na árvore de posições. O interpretador de *Ballcode* foi todo escrito em Javascript, contando com cerca de 600 linhas de código.

3.1.1 Leitura das linhas. Quando o usuário termina de escrever seu código em *Ballcode*, o primeiro passo do interpretador é a leitura do conteúdo bruto, que será dividido em linhas, que posteriormente serão alocadas em um vetor. Optou-se por utilizar o símbolo de ponto e vírgula (;) para delimitar o fim de uma linha. No caso das linhas que referenciam o início de blocos de código (linhas de *if* e *while*), o símbolo que delimita o fim da linha é o símbolo de dois pontos (:). Embora esses símbolos não sejam de fundamental importância para o interpretador, optou-se por deixá-los obrigatórios por fins pedagógicos. Os blocos de códigos são delimitados de forma semelhante aos blocos da linguagem Python: não são utilizadas chaves, e sim a indentação do código.

Durante a leitura do conteúdo bruto, cada símbolo de quebra de linha ('\n') adiciona o conteúdo lido até o momento no vetor, salvando assim cada linha do conteúdo bruto como uma *string* individual. Depois disso, tendo o conteúdo bruto dividido em linhas, é preciso normalizar este conteúdo.

3.2 Normalização das linhas

Tendo cada linha devidamente armazenada como *string*, é preciso classificá-las. Porém, antes desse processo, três coisas devem ser feitas: remover eventuais linhas em branco, descobrir se as linhas se encontram dentro de um bloco de código (pois isto será de fundamental importância quando gerarmos os elementos na tela) e remover eventuais redundâncias de espaçamento, tendo em vista que isso pode dificultar o processo de classificação.

Caso o conteúdo da *string* de linha seja vazio, ela simplesmente é excluída do processo, e nem mesmo é adicionada ao vetor de linhas em formato bruto. Também é importante salientar que os comandos, em hipótese alguma, podem ser quebrados em mais de uma linha.

Verificar se a linha está dentro de algum bloco de código é relativamente simples: como o conteúdo de cada linha foi obtido em sua totalidade, só precisamos verificar quantos símbolos de espaço ('`\u0020`') foram utilizados no começo da linha. Por motivos pedagógicos, a definição dos blocos é bastante restrita, e quaisquer erros de indentação podem (e muito provavelmente) gerarão erros de interpretação do código. Determinar se uma linha se encontra dentro de um bloco de código é um processo fundamental para o agrupamento dos objetos, tendo em vista que isso determinará a posição do objeto em relação aos objetos do contexto, já que eventuais erros de indentação fariam com que um objeto saísse da sequência desejada, ou mesmo passasse a compor outra estrutura (um comando que deveria ser do cano à direita do `if`, se indentado erroneamente, acabaria por não fazer parte do cano).

Após verificarmos se a linha pertence ou não a um bloco de código, removemos as redundâncias de espaçamento, substituindo qualquer quantidade de espaço por um espaço simples, de forma a facilitar o processo de classificação das linhas.

3.3 Classificação das linhas

Após a normalização das linhas, elas são classificadas. A classificação ocorre da seguinte maneira: cada linha é comparada com um conjunto de expressões regulares, sendo que cada expressão regular é referente a um tipo de linha diferente. Caso a comparação entre a linha e uma das expressões regulares retorne um valor verdadeiro, sabemos qual o tipo da linha, e esse dado será adicionado em um objeto do tipo linha. Caso a comparação com as expressões regulares não retorne verdadeiro em nenhum caso, podemos afirmar que a linha não constitui um comando válido, e isso retornará ao usuário um erro. Na Tabela 2, é possível ver as expressões regulares com as quais as linhas são comparadas.

3.4 Criação dos objetos de tipo Linha

Cada linha, depois de classificada, é enviada a um vetor de linhas, contendo os seguintes dados: tipo de linha, bloco e conteúdo. O valor do conteúdo possui os elementos de texto presentes na *string* bruta, tendo em vista que estes elementos serão utilizados para definir as características dos objetos, como o valor numérico das bolas, a operação dos canos, valores das estruturas de controle, etc.

3.5 Definição dos objetos em tela

Como o principal objetivo do jogo é demonstrar de forma visual o caminho pelo qual as variáveis do código passam, é de fundamental importância que o jogo seja assertivo neste ponto, definindo com

Table 2: Comandos em *Ballcode* e as Expressões Regulares referentes

Comando	Expressão Regular
Nova bola	<code>/^\s*new ball\s*=\s*\d{1,2}\s*;\s*\$/</code>
Atribuição	<code>/^\s*pipe is \[+/-*\]\s*\d{1,2}\s*;\s*\$/</code>
Estrutura Condicional	<code>/^\s*if ball\s*(== !=)\s*\d{1,2}\s*:\s*\$/</code> <code>/^\s*else\s*:\s*\$/</code> <code>/\s*endif\s*;\s*\$/</code>
Laço de Repetição	<code>/^\s*while ball\s*(== !=)\s*\d{1,2}\s*:\s*\$/</code> <code>/^\s*endwhile\s*;\s*\$/</code>

exatidão as posições dos objetos que estarão presentes em tela. Sendo assim, a ideia para definir as posições dos objetos foi o uso de uma árvore binária, já que a abstração usada para uma raiz e suas folhas se assemelha muito à abstração usada nos canos de estruturas de controle presentes no *Ballgorithm*. Esta foi a estrutura de dados utilizada para a organização dos elementos. Porém, com o passar do tempo, a estrutura mostrou-se limitada para lidar com os diferentes casos de canos. O fato das folhas das sub-árvores não guardarem quem são os seus nós-pais também se mostrou um problema, já que, para determinar as posições de certos elementos, é preciso visitar constantemente seus nós pais, e até mesmo avós. Desta forma, a árvore binária acabou sendo adaptada. A abstração que mais casava com a alocação dos canos era uma árvore ternária. Com a árvore ternária, é possível definir a posição dos objetos em tela apenas se baseando pela relação do objeto com o seu nó pai. Os canos sempre se localizarão abaixo de seu pai, e caso estes seja um filho à esquerda ou à direita, essa posição terá uma variação no eixo X.

Além disso, o último nó acessado na estrutura é guardado, evitando assim percorrer a árvore múltiplas vezes em busca de seu nó pai. Após a inserção dos objetos na estrutura, ela é percorrida, e durante esse processo os objetos têm suas posições na tela definidas, levando em consideração a sua relação com o nó raiz, que possui a posição inicial mapeada. Com isso, os processos que envolvem a interpretação das linhas terminam, e os objetos do jogo começam a ser criados.

3.6 Gerenciando os objetos em tela

Após a criação dos objetos na árvore, os elementos são alocados em vetores, agrupando os objetos de acordo com seus tipos. Tendo os grupos definidos, o interpretador manda um sinal para a *engine* gráfica, que começa a instanciar grupos de objetos, baseados nos vetores em que os objetos anteriormente foram alocados. Assim, cada objeto é instanciado com suas posições na tela e seus dados relevantes, como o símbolo da operação ou o tipo de comparação.

3.7 Instanciando os objetos

Os objetos de cada vetor são lidos, e deles são extraídos dados como suas respectivas posições em tela, informações sobre a operação referente ao objeto e valores. Os objetos recebem suas posições tendo a tela como base. A posição inicial no eixo X é a média entre 0 (zero) e o valor máximo, e cada elemento à esquerda deste objeto tem sua posição tirada da média entre o objeto inicial e 0 (zero), enquanto os elementos à direita tem suas posições definidas a partir da média entre o valor inicial e o valor máximo da tela, e assim os outros objetos tem suas posições determinadas recursivamente.

3.8 Definindo as características físicas dos objetos

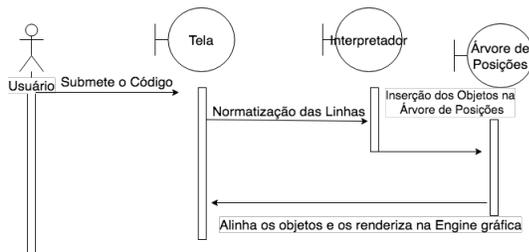
Depois de alocados os objetos em tela, é preciso determinar o que ocorre durante as colisões, já que cada elemento, quando em contato com as bolas, altera seus percursos ou valores. Assim, através da *engine* gráfica, podemos atribuir aos objetos regras de colisão e gravidade (no caso da bola). Por uma limitação da *engine* utilizada, não podemos aplicar colisões apenas nas paredes dos canos, o que torna necessário usar o seguinte truque: ao colidir com determinado cano, a posição da bola em (x, y) é mudada para a posição do cano posterior do percurso.

aprender a programar. Embora falhas de interface tenham sido apontadas nos questionários, estas falhas não foram determinantes para prejudicar de forma decisiva o uso da ferramenta. É importante também salientar que a ferramenta ainda pode oferecer tradução da linguagem para diversos idiomas, desafios entre os usuários e outros modos de jogo, que podem ser adequados de acordo com o tópico a ser ensinado.

REFERÊNCIAS

- [1] Rodrigo Pereira dos Santos and Heitor Augustus Xavier Costa. 2006. Análise de Metodologias e Ambientes de Ensino para Algoritmos, Estruturas de Dados e Programação aos iniciantes em Computação e Informática. *INFOCOMP 5*, 1 (2006), 41–50.
- [2] Dick Grune, Kees Reeuwijk, Henri Bal, Cerial Jacobs, and Koen Langendoen. 2012. *Modern Compiler Design*. Springer.
- [3] W. Huit. 2011. Motivation to Learn: An Overview. (2011). <http://www.edpsycinteractive.org/topics/motivation/motivate.html>
- [4] MEC/Inep. 2017. Taxa de conclusão média dos cursos de graduação presenciais. (2017). <http://www.observatoriodopne.org.br/metas-pne/12-ensino-superior/estrategias/12-3-fluxo/indicadores>
- [5] Luísa Behrens Palmeira and Matheus Parreiras Santos. 2015. Evasão no Bacharelado em Ciência da Computação da Universidade de Brasília: análise e mineração de dados. (2015).

Figure 6: Diagrama de seqüência do *Ballgorithm*



Após todos esses processos, os objetos são enfim renderizados em tela, e as bolas seguem o percurso automaticamente, fazendo os desvios definidos pelo usuário em seu código. Na Figura 6, é mostrada a ordem de execução de cada um dos passos citados anteriormente.

4 RESULTADOS

Testes preliminares da ferramenta foram feitos com usuários, e suas impressões foram captadas através de um Questionário de Satisfação do Usuário, abordando aspectos como facilidade de uso, didática da apresentação dos contextos, simplicidade do tutorial, além de outros aspectos funcionais, como velocidade de resposta e interface da ferramenta. Os resultados foram em grande maioria positivos no tocante à facilidade do uso da ferramenta e às abstrações utilizadas para os conceitos de programação.

5 CONCLUSÃO

Com este trabalho, pode-se concluir que o *Ballgorithm* é uma ferramenta multimídia bastante promissora, que pode ser utilizada em sala de aula e até mesmo por usuários finais com interesse em

WebMedia 2018
XXIV Simpósio Brasileiro de Sistemas Multimídia e Web

ANAIS

**V Workshop “O Futuro da
Videocolaboração”**

Coordenação

Leandro N. Ciuffo (RNP)

Valter Roesler (UFRGS)

Alexandre Carissimi (UFRGS)

V Workshop “O Futuro da Videocolaboração”

Sumário

Prefácio	158
Organização	162
Sessão Técnica	163
Performance de Música em Rede: Colaboração e prática musical à distância	163
Glayson Brendown Santos Silva (UFPB), Romero Coelho Alves (UFMG), Valdecir Becker (UFPB)	
Design audiovisual para experiência compartilhada: uma experiência de sincronização de conteúdo e interação aplicada à educação	167
Rafael Moura Toscano (UFPB), Durval Leal Filho (ONG Para’iwa), Marianna Cruz Teixeira (UFPB), Sandro Gonçalves (UFPB), Helder Bruno A.M. Souza (UFPB), Valdecir Becker (UFPB)	
VideoRecognition – Uma proposta de serviço para reconhecimento de elementos de vídeo em larga escala	171
Antonio J.G. Busson (PUC-Rio), Alan L.V. Guedes (PUC-Rio), Gabriel N.P. dos Santos (Faculdade ISL), Carlos de Salles Soares Neto (UFMA), Ruy Luiz Milidiú (PUC-Rio), Sergio Colcher (PUC-Rio)	
VideoViz – Uma proposta de ferramenta de visualização de logs de interação em players de vídeo	175
André Luiz de B. Damasceno (PUC-Rio), Alan L.V. Guedes (PUC-Rio), Sergio Colcher (PUC-Rio), Simone D.J. Barbosa (PUC-Rio), Dilson J.L. Rabêlo (UFMA), Elza B. Monier (UFMA)	
Mapping and review of audiovisual editors and streaming platforms based on collaborative technologies in real time	179
Brett Stalbaum (Univ. of California, US), Cicero I. da Silva (Unifesp), Jane de Almeida (Mackenzie Univ.), Maria Amelia Eliseo (Mackenzie Univ.), Vic Von Poser (Univ of the Arts, UK)	
Sistema de videocolaboração para o museu interativo do TCE-PB baseado no modelo DA	183
Amanda Azevedo (UFPB), Sandro Gonçalves (UFPB), Valdecir Becker (UFPB), Raoni Kulesza (UFPB), Fábila Carolino (TCE-PB), Agnaldo Macedo Filho (TCE-PB)	
Videoconferência multiambientes: o sistema Multipresença com foco na área da saúde ..	187
Valter Roesler (UFRGS), Guilherme Longoni (Mconf Tecnologia Ltda), Rafael Valle (RNP)	

V workshop “O Futuro da Videocolaboração”

Prefácio

Estabelecido em 2014 com o apoio a Rede Nacional de Ensino e Pesquisa (RNP), o Comitê Técnico de Videocolaboração (CT-Video) visa ser um fórum aberto para estudar o futuro das aplicações de vídeo e colaboração remota, nas mais variadas áreas, como por exemplo: Sistemas de videoconferência e telepresença; Sistemas de colaboração remota; Cinema digital; Aplicações multimídia na saúde; Produção, edição e transmissão multimídia; Áudio 3D; Educação online; Machine Learning em aplicações de vídeo; Software Defined Networking em aplicações de vídeo; Sistemas de vigilância e segurança; e Geração de vídeos em ambientes não controlados.

Qualquer pessoa pode participar do CT-Video. Seus membros são formados em sua maioria por: pesquisadores acadêmicos das áreas de multimídia, vídeo digital e redes; profissionais de empresas desenvolvedoras; usuários de serviços de videocolaboração. Para participar, basta se inscrever na lista de discussão (<http://listas.rnp.br/mailman/listinfo/ct-video>) do CT-Video e participar das reuniões. Adicionalmente, em 2018, foi estabelecido um grupo de pesquisa no CNPq com o nome “Videocolaboração”.

Um dos benefícios do Comitê é proporcionar a interação entre pesquisadores de diferentes instituições para facilitar a troca de conhecimento e eventual integração das aplicações desenvolvidas por cada grupo de pesquisa. Entretanto é importante destacar que o CT-Video está inserido em um projeto amplo de criação de comitês técnicos como forma de ampliar o processo de prospecção tecnológica da RNP. Além do CT-Video, outros dois comitês apoiam a RNP na realização da prospecção tecnológica: o comitê técnico de Monitoramento de Redes (CT-Mon) e o comitê técnico de Gestão de Identidade (CT-GId), sendo este último o mais antigo de todos, estabelecido em 2010.

O CT-Video realiza, em ciclos anuais, prospecção tecnológica com o auxílio de seus participantes, contribuindo para a construção de visões técnicas que orientem novos investimentos em Pesquisa e Desenvolvimento (P&D). Em geral, isso é materializado na forma de um mapeamento sistemático do estado da arte das pesquisas em andamento nos diversos laboratórios participantes do CT-Video e na elaboração de um relatório anual de Visão de Futuro.

Os 7 artigos aprovados para o V workshop do CT-Video foram avaliados pelo comitê de programa

do evento, que forneceram sugestões que foram consideradas para a elaboração das versões finais publicadas nestes anais.

Laboratórios e Grupos de Pesquisa que contribuíram para este workshop

Apresentados na mesma ordem de autoria dos artigos:

O **Grupo de Pesquisa Design Audiovisual da UFPB** foi criado em 2016 e é composto por professores, doutorandos, mestrandos, graduandos e alunos de iniciação científica de diferentes cursos e universidades, Seu foco está na integração metodológica e teórica dos processos de criação, produção, desenvolvimento e fruição das áreas da Interação Humano Computador com os estudos de audiência, que tem metodologias e propostas similares, apesar de estarem em áreas do conhecimento distintas. Desenvolve pesquisas e projetos com temas variados, tais como: narrativas generativas, conteúdos pedagógicos, jornalismo 360o, visualização de dados, conteúdos para TVs públicas, rádio e interatividade.

O **TeleMídia** é um laboratório de pesquisa afiliado ao Departamento de Informática da PUC-Rio, dedicado à pesquisa e inovação em sistemas hipermídia e multimídia distribuídos. O TeleMídia foi fundado em 1991 pelo Prof. Luiz Fernando Gomes Soares, e é atualmente liderado pelo Prof. Sérgio Colcher. Além da sua vasta contribuição em pesquisas de base e em formação de recursos humanos, nos últimos anos, o laboratório também se destacou com importantes contribuições relacionadas à padronização em sistema de TV digital terrestre e IPTV. Entre as mais conhecidas dessas contribuições estão a linguagem NCL (Nested Context Language) e o middleware Ginga, ambos padrões para serviços interativos no Sistema Brasileiro de TV Digital Terrestre e recomendações da União Internacional das Telecomunicações para IPTV e sistemas híbridos broadcast/broadband.

O **Núcleo Técnico-Científico (NTC) Telessaúde Brasil Redes Unifesp** é uma parceria entre a Universidade Federal de São Paulo (Unifesp) e o Ministério da Saúde. A parceria é estabelecida por meio de convênio, em regime de cooperação técnica, administrativa e científica, com o objetivo de manter o Núcleo na Unifesp atuante e com cobertura no Estado de São Paulo, prioritariamente nos municípios de difícil acesso e áreas remotas de maior vulnerabilidade, para melhoria da qualidade da assistência na Atenção Básica/Atenção Primária em Saúde no Sistema Único de Saúde, utilizando diferentes estratégias educacionais, formativas e de suporte. O NTC provê acesso a mais de 1000 equipes de Saúde da Família, com atividades que vão desde teleassistências e telediagnósticos por meio de recursos digitais avançados (vídeo ou webconferências síncronas e assíncronas), telediagnósticos com o uso de aplicações avançadas de visualização remota, ações de formação

educacional em APS (formação permanente dos profissionais de saúde) e o desenvolvimento de Segundas Opiniões Formativas (SOF's) para apoio à pesquisa e educação permanente a partir da publicação nas bibliotecas virtuais de saúde (BVS's) e bases indexadas nacionais e internacionais. O NTC tem como membros os pesquisadores do Grupo de Pesquisa em Saúde Digital: Telessaúde e Tecnologias da Informação: <http://dgp.cnpq.br/dgp/espelhogrupo/1179489123867343> .

O Laboratório de Artes Cinemáticas e Visualização (LabCine) da Universidade Presbiteriana Mackenzie produz pesquisas históricas, teóricas, críticas e tecnológicas no campo das artes audiovisuais, a partir de colaboração com pesquisadores de tecnologias computacionais e de redes. O LabCine desenvolve e pesquisa cinemas expandidos, modos de exibição em espaços contemporâneos, arquivamento e seleção de imagens, narrativas científicas e educacionais, visualizações avançadas de bancos de dados e visualizações avançadas remotas. O LabCine foi criado para dar continuidade a pesquisas em um campo que é mais amplo que o cinema tradicional e por isso a denominação “cinemáticas”. O grupo, registrado no CNPq, tem mais de 20 participantes entre doutores, graduandos, mestrandos, doutorandos, pós-doutorando e é composto por pesquisadores de diferentes universidades como UNICAMP, UFG, UNIFESP e PUC/SP e parceiros de pesquisa da UCSD (Universidade da Califórnia de San Diego), UFPB e USP.

O Laboratório de Aplicações de Vídeo Digital (LAVID), integrado ao Departamento de Informática (DI) da Universidade Federal da Paraíba (UFPB), foi criado em 2003. O laboratório surgiu da proposta de desenvolver projetos de pesquisa em hardware e software voltados às áreas de Vídeo Digital, Redes de Computadores, TV Digital e Interativa e Middleware. Atualmente o LAVID é uma referência nacional e internacional em desenvolvimento de tecnologia para TV Digital. O LAVID conta com a colaboração de mais de 40 jovens pesquisadores, entre doutores, mestres e graduandos, que estão interconectados com pesquisadores de todo o Brasil e do mundo, trazendo as atuais tendências tecnológicas mundiais nas áreas de vídeo e TV Digital.

O Laboratório do PRAV (Projetos em Áudio e Vídeo) da UFRGS iniciou suas atividades em 1998 e, desde 2006, está situado junto ao Instituto de Informática da Universidade Federal do Rio Grande do Sul, sendo composto atualmente por quatro áreas de especialidade: Desenvolvimento de sistemas para videoconferência; Telemedicina / Telessaúde; Ensino a Distância; Redes de Computadores. Com uma equipe de mais de 30 pesquisadores e colaboradores, o PRAV conquistou uma posição de referência em tecnologia no Brasil, tendo diversas parcerias com outras Universidades, grupos de pesquisa e empresas. O laboratório do PRAV foi berço de duas spin-offs. A empresa I9Access

Tecnologia Ltda trabalha na área de Sistemas de Telessaúde (Sala Cirúrgica Integrada, Sistemas de simulação realística assistida por vídeo, Sistemas de monitoramento de doenças crônicas, tele-ECG, entre outros). A empresa Mconf Tecnologia Ltda tem seu foco no sistema de webconferência Mconf, voltado para reuniões remotas e ensino a distância.

Agradecemos a todos os autores que submeteram seus trabalhos.

Agradecemos ao Professor Manoel C.M. Neto (IFBA), pelo seu intenso trabalho e dedicação na organização do WebMedia 2018, e ao Professor Roberto Willrich (UFSC) pela cuidadosa edição destes anais.

Agradecemos também à CE-WebMedia pelo respaldo ao workshop do CT-Video e pela confiança depositada no nosso trabalho.

Salvador, outubro de 2018.

Leandro N. Ciuffo (RNP)

Valter Roesler (UFRGS)

Alexandre Carissimi (UFRGS)

Coordenadores do V Workshop “O Futuro da Videocolaboração” do WebMedia 2018

Coordenadores do V WCT-VIDEO

Leandro N. Ciuffo (RNP)
Valter Roesler (UFRGS)
Alexandre Carissimi (UFRGS)

Comitê de Programa do V WCT-VIDEO

Alexandre Carissimi (UFRGS)
Leandro N. Ciuffo (RNP)
Paulo Lopes (RUTE)
Valter Roesler (UFRGS)

Performance de Música em Rede

Colaboração e prática musical à distância

Glayson Brendown Santos Silva

Programa de Pós-Graduação em
Computação, Comunicação e Artes,
Centro de Informática, UFPB.
glayson.silva@lavid.ufpb.br

Romero Coelho Alves

Arte e Mídia, Centro de Humanidades,
Universidade Federal de Campina
Grande, UFCG.
rcoelhoalves@gmail.com

Valdecir Becker

Programa de Pós-Graduação em
Computação, Comunicação e Artes,
Centro de Informática, UFPB.
valdecir@ci.ufpb.br

ABSTRACT

This theoretical and experimental work provides an introductory view about the universe of network music performance (NMP). Through collaboration between musicians and researchers from the Universidade Federal da Paraíba and the Universidade Federal de Campina Grande, we performed an NMP experiment and from it we verified and analyzed the advantages, disadvantages and artistic and creative possibilities inherent in the use of the NMP ReaNINJAM tool.

KEYWORDS

Network music performance, NMP

1 INTRODUÇÃO

O potencial na performance de música em rede (*network music performance* - NMP) é expressivo quando tratamos de aplicações na Internet, pois além de fornecer a possibilidade de músicos remotos tocarem juntos, propende a uma revolução no conceito tradicional de interação musical [6]. Com a evolução da tecnologia, a NMP vem sendo experimentada, mas ainda enfrenta diversos desafios. Alguns desses empecilhos são inerentes às características da Internet/rede de computadores. Dentre eles podemos citar a latência, falta de sincronia audiovisual e sensibilidade à perda de informação na rede [1].

As primeiras experiências de colaboração musical remota datam da época de John Cage com sua obra *Imaginary Landscape No. 4 for twelve radios*, composta em 1951 [5]. Outro exemplo é a *League of Automatic Music Composers*, uma banda/coletivo californiana de música eletrônica experimental que atuou entre 1977 e 1983, que executava músicas utilizando uma rede local de computadores [2, 4]. Essas experiências são indícios do que estaria por vir com a evolução da tecnologia nas áreas de hardware e software de computadores, Internet e transmissão de dados na rede, dentre outros avanços que permitiram que esses aspectos se

aproximassem cada vez mais de nós, facilitando nossas vidas e proporcionando novas experiências.

A fim de refletir os possíveis impactos da tecnologia de NMP, bem como analisar as possibilidades artísticas que com ela surgem, fizemos uso de um *software* de computador que possibilita pessoas de diferentes cidades, estados ou países, tocarem juntas como se estivessem reunidas num mesmo lugar. Através da colaboração entre pesquisadores de duas instituições, sendo uma delas a Universidade Federal da Paraíba (UFPB) e a outra a Universidade Federal de Campina Grande (UFCG), experienciamos a NMP para obter diferentes perspectivas do uso dessa ferramenta.

2 TRABALHOS RELACIONADOS

Existem alguns trabalhos que tratam do panorama geral da NMP, dentre eles podemos citar o *Networked Music Performance: State of the Art* [5], que aborda o estado da arte do NMP, bem como o *An Overview on Networked Music Performance Technologies* [6]. No levantamento realizado por esses estudos, é notória a prioridade dada a fatores mais técnicos como soluções para latência, arquiteturas de *software* desenvolvidas para as plataformas, aprimoramento de algoritmos, e assim por diante, mas a preocupação com o tópico de estudo proposto por nós, que visa analisar o potencial artístico e criativo, bem como refletir acerca das possibilidades que a NMP oferece, não se apresenta.

3 PERFORMANCE DE MÚSICA EM REDE

A performance de música em rede acontece quando músicos geograficamente dispersos tocam em conjunto através de uma rede de computadores conectados entre si, como por exemplo, através da Internet [1, 5, 6]. Segundo Carot e Werner [3], podemos classificar a NMP em 4 principais abordagens: interação realista (*realistic interaction approach* – RIA), mestre-escravo (*master slave approach* – MSA), descontraída (*laid back approach* – LBA) e de feedback retardado (*delayed feedback approach* – DFA).

A RIA é uma interação musical realista que implica numa latência unidirecional estável, menor que 25 milissegundos, entre os dois instrumentos responsáveis pelo ritmo, como uma bateria e um contrabaixo. Segundo Carot e Werner [3], se imaginarmos dois

Em: V Workshop “O Futuro da Videocolaboração” (WCT-Video 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pósteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.
©2018 SBC – Sociedade Brasileira de Computação.
ISBN: 978-85-7669-435-9.

músicos interagindo juntos num palco, a uma distância de 8 metros, o tempo de reação entre o que um que toca e o que o outro ouve pode ser calculado. Dado que o som viaja a 340 m/s através do ar, o tempo de latência entre os músicos é de aproximadamente 25 milissegundos. Em razão disso, numa orquestra, por exemplo, seria inviável a execução de uma peça sem a figura do maestro para conduzir o tempo, pois a disposição dos músicos pode ultrapassar os 8 metros, o que ocasiona a falta de sincronismo por conta da distância entre eles. A RIA é o único cenário no qual o músico pode tocar como se estivesse num ensaio, compondo, ou num show, pois quando a latência supera os 25 ms, o processo de construção musical começa a depender de outros fatores inerentes a essa latência.

Na MSA existe uma relação de mestre e escravo entre os instrumentos rítmicos. Os músicos devem compreender que a perfeição e a sincronia musical serão diferentes em relação a execuções presenciais, uma vez que nessa abordagem eles lidam com um atraso maior que 25 ms. Sendo assim, um músico fica responsável por comandar o *groove* (mestre), enquanto o outro apenas o segue (escravo). Para o mestre não haverá uma sincronização satisfatória, por isso ele norteia a música. Por outro lado, o escravo tem total sincronismo entre os dados que recebe e os que envia, sem atraso algum.

Se substituirmos um instrumento rítmico por um solo, por exemplo, trocar a bateria por uma flauta, temos como abordagem satisfatória a LBA. Nesse contexto, o atraso gira em torno de 50 a 100 ms, mas não atrapalha a experiência musical. A LBA é parecida com a MSA, pois o músico responsável pelo instrumento rítmico comanda o *groove*, enquanto o solista fraseia em cima do que está sendo proposto. Para o solista não há atraso e a sincronia é perfeita, pois quem trabalha com a latência é o músico rítmico. Por isso este deve considerar que o atraso no recebimento dos dados faz parte da experiência musical e compõe a atividade. Como essa abordagem propõe uma improvisação mais permissiva, como no jazz, a latência não é um problema, pois nesse contexto o objetivo não é soar uníssono, tendo as duas partes que se encaixarem perfeitamente na mesma batida e tempo.

Por fim, a DFA consiste em atrasar o sinal de entrada, de ambos os músicos, para gerar a sincronia. Essa abordagem tenta melhorar a experiência do mestre, que na MSA e na LBA conta com atraso no recebimento de dados do músico remoto. Para este fim, a DFA atrasa o sinal de entrada do músico local e do remoto, para que os dois possam ouvir o que estão tocando numa sincronia mais satisfatória. Caso o atraso aconteça, devido a alguma instabilidade na conexão, por exemplo, a experiência pode gerar incômodos, atrapalhando os músicos, que ouvem seu sinal de entrada alguns milissegundos depois do que está executando de fato.

De acordo com Carot e Werner [3], além dessas 4 principais abordagens que pretendem lidar com a latência, para que a musicalidade seja a mais realista possível, existe uma outra que não tenta solucionar os problemas de atraso. A abordagem de aceitação de latência (*latency accepting approach* – LAA), como o próprio nome sugere, aceita atrasos acima de 25 ms, pois o objetivo não é

obter músicas convencionais, mas sim a livre experimentação. Neste contexto os músicos aceitam a latência e a utilizam como forma de expressão artística, o que faz da LAA uma abordagem vanguardista.

Visando fornecer uma visão da NMP que possa ser alcançada pelo maior número de pessoas, optamos por utilizar recursos como computadores comuns e Internet comercial, cuja banda larga é de 25 megabytes. Em razão disso, a abordagem escolhida para a experiência foi a DFA.

4 EXPERIMENTANDO A NMP

Com o objetivo de dissertar sobre uma experiência de performance de música em rede alcançável pelo maior número de pessoas possível, decidimos utilizar o *software* Reaper¹ (Fig 1.), que oferece dois meses de uso gratuito. Esse *software*, que é uma estação de trabalho de áudio digital voltada para gravação e mixagem, conta com o plugin nativo ReaNINJAM, o qual deve ser utilizado para realizar a NMP.



Figura 1: Interface do software Reaper

O experimento que realizamos contou com a colaboração de três músicos localizados em posições geográficas distintas no estado da Paraíba, estando dois na cidade de Campina Grande, em residências diferentes, e um em João Pessoa.

O intuito do experimento foi explorar a NMP, fazendo uma comparação com a experiência tradicional de tocar em conjunto quando os músicos estão no mesmo lugar, para assim perceber quais novidades esse tipo de performance oferece, sejam positivas ou negativas.

O ReaNINJAM oferece algumas salas (Fig. 2) nas quais os participantes podem se conectar para participar de uma *jam*², mesmo estando em cidades, estados ou países diferentes. Após selecionada a sala, o participante está pronto para tocar com os outros (Fig. 3).

O processo aconteceu em três sessões de 50 minutos cada, realizadas em dias diferentes, para uma melhor absorção dessa nova maneira de tocar com o outro e adaptação a esse meio. Os instrumentos selecionados para o experimento foram guitarra, contrabaixo e teclado. Na primeira sessão, os três músicos se

¹ <https://www.reaper.fm/>

² Termo comumente utilizado por músicos quando se referem a tocar sem compromisso, improvisando, sem ensaiar.

reuniram sozinhos numa sala para realizarem uma *jam*, no intuito de conhecer o *software* e entender a forma como ele funciona. Na segunda sessão, os músicos se reuniram sozinhos mais uma vez, porém buscando tocar alguma música comum ao conhecimento dos três, para checar a influência que a latência causa na sincronia do grupo. Na terceira sessão, os músicos entraram numa sala com outros participantes, buscando experimentar uma *jam* com músicos que possivelmente seriam de diferentes cidades, estados ou países.

Na sala de *chat* (Fig. 4) podemos observar um endereço IP após o nome que identifica cada participante. Através do uso de um localizador de IP³, verificamos que existem endereços associados a diferentes países, como África do Sul, Alemanha e Malásia.

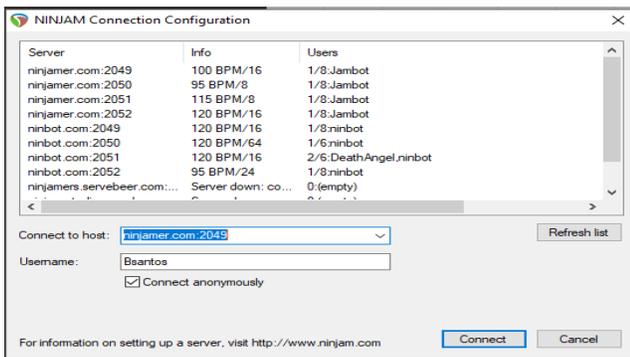


Figura 2: Salas de NMP do ReaNINJAM

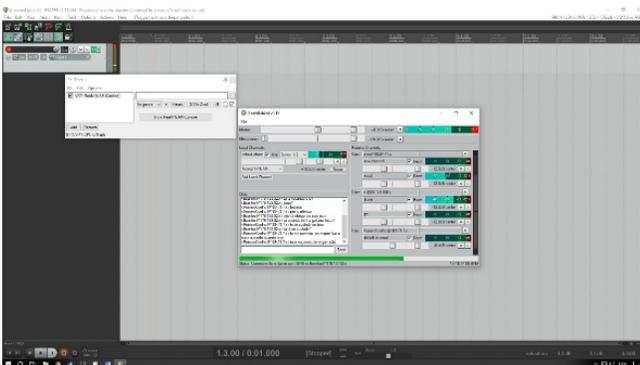


Figura 3: Captura de tela durante jam no ReaNINJAM

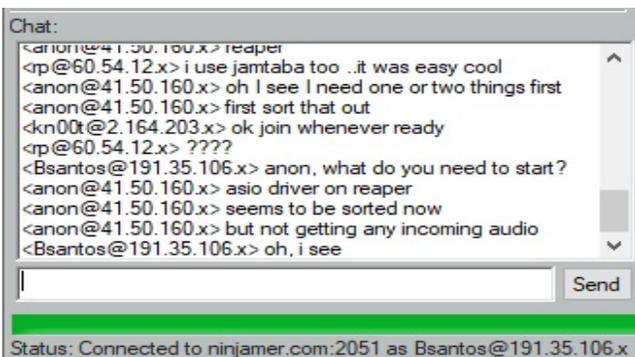


Figura 4: Sala de chat do ReaNINJAM

5 RESULTADO E DISCUSSÃO

Acreditamos que acessibilidade e aproximação estão diretamente ligadas ao que a Internet representa, e a experiência com a performance de música em rede nos provou isso. Ao oferecer a possibilidade de músicos dispersos geograficamente tocarem em conjunto, a NMP tanto pôde aproximar culturas como facilitar o diálogo entre elas. A prática musical antes realizada num palco, num estúdio de gravação ou mesmo numa garagem, agora pode ser realizada frente à tela de um computador. A necessidade da presença física do outro para tocar em conjunto não se aplica mais da mesma forma, o que dá espaço para novos tipos de performance surgirem.

As três sessões de performance de música em rede foram realizadas com sucesso. Os músicos, apesar de na primeira sessão estarem desconfortáveis com os aspectos técnicos da ferramenta, apreciaram a NMP ao passo que puderam criar intimidade com o processo. Depois de se acostumarem à latência, a musicalidade não demorou a fluir e logo os três participantes puderam ter uma experiência satisfatória. Já na segunda sessão, a tentativa de executar uma música comum aos três foi inicialmente um fracasso. Como na música escolhida os três instrumentos deveriam iniciar ao mesmo tempo, a rítmica da execução foi totalmente comprometida, pois a latência impediu que houvesse sincronismo entre as partes. Em razão disso, os músicos decidiram improvisar em cima dessa mesma música, mas dessa vez o teclado serviu como guia, para que a guitarra e o contrabaixo apenas o seguissem. Na terceira sessão, ao entrar na sala, os músicos se depararam com uma *jam* que já tinha sido iniciada, portanto seguiram as ideias musicais e colaboraram com seus repertórios.

O experimento resultante desse trabalho revelou, através da análise das opiniões dos músicos que participaram, diversos aspectos que são frutos da NMP, dentre eles as vantagens, desvantagens e possibilidades que esse tipo de performance musical nos oferece.

5.1 Vantagens da NMP

Acreditamos que dentre as vantagens que a performance de música em rede nos oferece, a mais básica delas é a de poder tocar com outro músico a qualquer hora sem haver necessidade de sair de casa, desde de que haja participantes online. Como a experiência na NMP acontece através de um computador, o músico pode optar por utilizar fones de ouvido para monitorar sua performance, gerando baixíssimo ruído a depender do instrumento que está executando. Essa condição fornece a possibilidade de tocar em qualquer horário do dia (novamente, a depender do instrumento que esteja sendo executado) sem haver a necessidade de ter um ambiente controlado para conter as ondas sonoras geradas, por exemplo, por um amplificador de guitarra ou por monitores de áudio. Logo, o participante que mora em casa ou apartamento, pode desfrutar de uma *jam* sem ter maiores preocupações.

Para aqueles que estudam música, a NMP também oferece a vantagem de tocar com outras pessoas, que diferentemente das *backing tracks*, utilizadas por instrumentistas que gostam de

³ <https://www.iplocation.net/>

improvisar, reagem de forma orgânica ao que ouvem, acarretando numa experiência de estudo mais próxima de performances de música ao vivo. Enquanto a *backing track*, que serve de acompanhamento, é repetitiva e previsível, o participante da NMP pode ofertar novidades a cada novo compasso da *jam*.

A NMP proporciona a aproximação musical do outro, em razão disso não são raras as *jams* que contenham participantes de diversas cidades, estados ou países diferentes. Nosso experimento, por exemplo, possivelmente abrangeu 4 continentes diferentes, como a América do Sul, África, Ásia e Europa. Essa característica resulta numa mixagem de culturas diversas, que através da música se unem e constroem ricas sonoridades, o que é de grande valia ao músico, pois este está constantemente agregando repertório ao seu conhecimento musical, que enriquece através desse contato com diferentes formas de pensar música.

Acreditamos que, graças a aproximação cultural que pode existir na NMP, a vantagem mais rica é a de poder compor músicas à distância, e mais do que isso, com participantes de diferentes países. Como a linguagem musical é universal e independe da língua falada ou escrita, novas formas de comunicação e de arte surgem, revelando vieses criativos inéditos. Colaborar com culturas diferentes em prol da produção artística musical é um forte aspecto positivo que a NMP proporciona, pois democratiza o processo de composição musical e pode torná-lo cada vez mais diverso, o que acarreta numa variedade maior de possibilidades de surgimento de novos gêneros musicais.

5.2 Desvantagens da NMP

Das desvantagens da NMP, consideramos a mais básica como sendo a necessidade de o participante ter algum conhecimento técnico na área de informática. Para utilizar a ferramenta é necessário lidar com *downloads* e instalações de *softwares* e *plugins*. Apesar das exigências não serem significativas, a instalação e configuração do *software* representa um fator limitante para parte dos músicos.

Outra característica que implica na experiência de NMP é a adaptação a qual o músico deve passar para utilizar de forma fluida a ferramenta, tendo em vista que é um ambiente diferente do qual normalmente está acostumado a tocar. A falta de intimidade com as ferramentas de NMP pode implicar um bloqueio criativo ou uma insatisfação que gere uma má execução do instrumento, o que acarreta uma performance desagradável.

Finalmente, o principal aspecto negativo que encontramos no experimento foi a latência. Salve algumas exceções, como grandes orquestras ou shows em grandes arenas e palcos, os músicos não lidam com latência, logo enfrentar atrasos é a principal dificuldade a ser encarada no processo. Essa limitação da NMP é intrínseca, já que a ferramenta utiliza a Internet para funcionar. Trata-se, portanto, ainda de um problema estrutural, sem solução a curto prazo. Dessa forma, é pertinente que os músicos incorporem a latência ao processo criativo e de execução.

Dada a limitação física da latência, preferencialmente a NMP é utilizada para a realização de *jams*, pois estas priorizam a improvisação, o que pode afetar o interesse daqueles que buscam

tocar músicas finalizadas que necessitam serem executadas num só tempo, respeitando a cronologia dos compassos.

5.3 Possibilidades da NMP

Analisando as vantagens e desvantagens da NMP, refletimos acerca das possibilidades que podem surgir através da experiência. Surge, com a performance em rede, um novo meio de compor música, não tradicional, que apresenta dinâmicas e nuances diferentes, portanto gera resultados diferenciados. O principal aspecto não tradicional para essa nova forma de fazer música é considerar a latência como elemento musical e não como ruído. Utilizar o atraso como acessório para a composição é expandir o leque de opções de construção de uma música.

Para uma nova forma de composição musical, surge uma nova forma de apreciar esse tipo de música. Não só os compositores deverão se reeducar para compor peças que insiram a latência como parte da música, mas também os ouvintes precisarão obter uma nova postura para consumir essas sonoridades não tradicionais. Trata-se de uma expressão artística inerente ao universo computacional e em rede.

Por fim, acreditamos que essa possibilidade da inserção proposital da latência na música afeta diretamente a estética que rege esse ramo da arte, pois o que é considerado ruído pode tornar-se adereço estético, depositado na composição com a intenção de comunicar.

6 CONCLUSÃO

A realização desse estudo teórico e experimental proporcionou uma visão acerca do universo da performance de música em rede. Através do experimento realizado com a colaboração de três músicos, pudemos verificar e analisar as vantagens, desvantagens e as possibilidades fornecidas pela ferramenta de NMP ReaNINJAM. Acreditamos que a NMP será um método utilizado com mais frequência futuramente, pois a evolução e democratização tecnológica permite acesso a recursos cada vez mais avançados, como Internet mais veloz e estável, e computadores mais potentes, o que afeta diretamente o funcionamento dos *softwares* da performance de música em rede. Esperamos que esse trabalho sirva de repertório para pesquisadores da área, bem como de introdução aos interessados pelo assunto.

REFERÊNCIAS

- [1] Akoumianakis, D. et al. 2014. *The MusiNet project: Towards unraveling the full potential of Networked Music Performance systems*. IISA 2014 - 5th International Conference on Information, Intelligence, Systems and Applications.
- [2] Barbosa, A. 2003. *Displaced soundscapes: A survey of network systems for music and sonic art creation*. Leonardo Music Journal, v. 13, p. 53–59.
- [3] Carot, A.; Werner, C. 2009. *Fundamentals and Principles of Musical telepresence*. Journal of Science and Technology of the Arts, v. 1, n. 1, p. 26.
- [4] Follmer, G. 2005. *Electronic, aesthetic and social factors in net music*. Organised Sound, v. 10, n. 3, p. 185–192.
- [5] Renaud, A.; Carot, A.; Rebelo, P. 2007. *Networked Music Performance: State of the Art*. Media, n. 4, p. 1–7.
- [6] Rottondi, C. et al. 2016. *An Overview on Networked Music Performance Technologies*. IEEE Access, v. 4, p. 8823–8843.

Design audiovisual para experiência compartilhada:

uma experiência de sincronização de conteúdo e interação aplicada à educação

Rafael Moura Toscano
GP - Design Audiovisual
João Pessoa, Paraíba.
rafaeltoscano@lavid.ufpb.br

Durval Leal Filho
ONG Para'íwa
João Pessoa, Paraíba.
durvalleal@gmail.com

Marianna Cruz Teixeira
GP - Design Audiovisual
João Pessoa, Paraíba.
mariannactx@gmail.com

Sandro Gonçalves
GP - Design Audiovisual
João Pessoa, Paraíba.
sandro.goncalves@lavid.ufpb.br

Helder Bruno A. M. Souza
GP - Design Audiovisual
João Pessoa, Paraíba.
brunoarpeggi@gmail.com

Valdecir Becker
GP - Design Audiovisual
João Pessoa, Paraíba.
valdecir@ci.ufpb.br

ABSTRACT

The development of audiovisual content has become an increasingly complex activity. In view of this scenario arises the Architecture of Audiovisual Design (AD) as a theoretical-methodological instrument of analysis and creation of systems. The purpose of this article is to describe the process of building the multi-screen audiovisual system Caatinga Viva from AD. To achieve this goal, a partnership was established between Para'íwa and the Audiovisual Design Research Group to create an experiment in education. The construction of the system revealed that the elements available in the DA as Design Lines, roles and triggers are potential elements to create multi-screen systems.

KEYWORDS

Audiovisual Design, multi-screen, interaction, video collaboration

1 INTRODUÇÃO

O desenvolvimento de conteúdos audiovisuais tem se tornado uma atividade cada vez mais complexa a medida que novos recursos de interação, compartilhamento, diferentes telas e formas de visualização vem se popularizando. A experiência audiovisual, que antes era majoritariamente ocupada pelo cinema e televisão, hoje torna-se mais abrangente, podendo envolver desde pesquisas na Internet a complexos algoritmos de recomendação, colaboração e adaptação de conteúdos. Uma característica deste cenário é a presença e o uso de softwares para compor a experiência de fruição das obras audiovisuais.

Em paralelo as discussões da Computação, nos estudos de mídia diferentes modelos foram sugeridos ao longo dos anos com o objetivo de descrever o processo de transmissão e recepção de conteúdos [13]. Recentemente o estudo [6] propôs três modelos gerais para descrever diferentes cenários: a de um emissor para muitos receptores, como Broadcast; a comunicação online em que o indivíduo toma a iniciativa de buscar o conteúdo (Stickiness); e Propagável (Spreadable), quando o conteúdo alcança a audiência

por meio da ação de outros indivíduos, a partir de ferramentas essencialmente digitais. Segundo os autores, estes modelos coexistem com proporções e intensidades distintas. Apesar de não descreverem completamente o estado da arte da utilização de mídia, os modelos Broadcast, Stickiness e Spreadable nos ajudam a perceber a crescente presença de processos midiáticos no cotidiano das pessoas, bem como a necessidade de criação de novas estratégias que comportam cenários em que a atenção é partilhada por processos e meios distintos.

Diante deste contexto complexo da utilização de mídias audiovisuais, surge a discussão teórica-metodológica do Design Audiovisual (DA). Inicialmente proposta no artigo [1], o DA integra aspectos de Mídia e Recepção da Comunicação a ferramentas e conceitos da Interação Humano Computador (IHC), com o intuito de auxiliar na descrição, análise e criação de sistemas audiovisuais. Em uma segunda publicação o DA é caracterizado pelos autores como uma arquitetura de conteúdo e interação capaz de atender a contextos de videocolaboração [14]. Com o intuito de avançar a proposição dos estudos citados, este trabalho propõem a aplicação experimental do Design Audiovisual como a arquitetura de um projeto de videocolaboração multi-tela. Tal experimento foi idealizado e executado pela Organização Não Governamental Para'íwa em colaboração com pesquisadores do Grupo de Pesquisa Design Audiovisual (GPDA).

A ONG Para'íwa¹ atua desde 1994 no estado da Paraíba, com ações direcionadas ao registro e promoção da cultura, diversidade, meio ambiente, economia criativa e práticas cidadãs através de projetos relacionados as mais diversas linguagens e Tecnologias da Comunicação e Informação (TIC's). Já o GPDA foi criado em 2016 e conta com professores, doutorandos, mestrands e graduandos de diferentes cursos e universidades do Brasil. O foco do grupo está na construção teórica-metodológica do Design Audiovisual a partir de contextos como: mídias generativas, rádio, objetos educacionais, jornalismo 360°, visualização de dados e interatividade.

Como resultado da colaboração técnico-científica de pesquisadores e profissionais destas instituições temos a aplicação do Design Audiovisual como fundamento teórico-metodológico de um sistema de videocolaboração multi-tela. A construção deste artefato, denominado de Caatinga Viva, tem o objetivo de aplicar a arquitetura

In: V Workshop "O Futuro da Videocolaboração" (CT-Vídeo 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pósteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.

© 2018 SBC – Sociedade Brasileira de Computação.
ISBN 978-85-7669-435-9.

¹Disponível em: <<http://www.paraíwa.org.br>>

apresentada nos estudos [1, 14] em um contexto de interação síncrona entre múltiplos conteúdos e dispositivos. Todo este processo ocorre tecnicamente a partir de um ambiente de rede bidirecional e persistente que permite a criação de novas interfaces de visualização.

As demais seções do trabalho detalham a pesquisa da seguinte forma: a segunda seção descreve discussões e sistemas que dialogam com o escopo deste estudo; a terceira seção delimita os conceitos essenciais para a compreensão da arquitetura do Design Audiovisual (DA); na quarta seção são descritas as etapas da pesquisa e a colaboração dos participantes; a quinta seção descreve a arquitetura de interação do sistema. Por fim, apresenta-se as lições aprendidas e perspectivas futuras do estudo.

2 TRABALHOS RELACIONADOS

O desenvolvimento de sistemas audiovisuais com essa natureza complexa pode ser identificado em diversos produtos e serviços disponíveis atualmente. Aplicações multi-tela para TV digital, Smart TV's, centrais de mídia tipo Apple TV ou Amazon Fire e WebApps como o sistema Air Console ² estão cada vez mais populares e acessíveis ao grande público. Apesar da popularização destes sistemas são poucos os estudos a nível teórico-metodológico que fundamentam a criação de conteúdo e interação para este novo ambiente de mídia, uma ambiência de colaboração em vídeo por parte de sistemas e usuários.

A busca pela construção de uma mídia audiovisual com suas funções expandidas em camadas de software pode ser identificada em estudos mais antigos como na pesquisa [12]. Esta expansão da experiência de recepção ao áudio e vídeo desenvolveu-se ao longo dos anos basicamente em duas perspectivas: a multimedia e o hypervídeo [8]. Especificamente sobre a relação de sincronia entre dois dispositivos, costuma-se usar o termo Segunda Tela. As discussões sobre experiências deste tipo concentram-se majoritariamente nas questões de sincronismo, todavia estudos como o [4] destacam recomendações gerais para o desenvolvimento do conteúdo. Em um nível maior de abstração é possível encontrar estudos como [2, 3, 5, 7, 9–11] que descrevem noções básicas de engajamento ou convenções para o uso de mídia. Apesar de paralelamente autores da computação e comunicação apresentarem relatos sobre aplicações multi-tela não há um escopo teórico que integre estes conhecimentos. Uma discussão potencial ao campo de aplicações multi-tela é a arquitetura do Design Audiovisual, uma vez que considera como parte essencial da fruição as relações que um indivíduo desenvolve tanto com o conteúdo quanto com o sistema.

3 FUNDAMENTAÇÃO TEÓRICA

A arquitetura de design colaborativo desenvolvida nos estudos [1, 14] descreve quatro funções, ou papéis, que os indivíduos podem assumir enquanto consomem conteúdo audiovisual: Audiência, Sintetizador, Modificador e Produtor. Cada papel tem seus níveis elevados, descritos como Player, quando os indivíduos se envolvem e usam todos os recursos disponíveis, ou, ainda, subvertem os propósitos do conteúdo e do sistema.

Para o Design Audiovisual as relações de interação inerentes a esses níveis ocorrem por meio de quatro Linhas de Design: (a)

Conteúdo; (b) Identidade, que baseia a relação pessoal com o conteúdo; (c) Motivação, que leva ao engajamento; e (d) Experiência, que relaciona a fruição à utilização avançada de recursos, que devem proporcionar uma experiência prática ou simbólica com o conteúdo e interfaces. As Linhas de Design tem o intuito de auxiliar o desenvolvedor do sistema audiovisual (conteúdo e software) na elaboração do projeto de interação, engajamento e conteúdo de sua produção a partir de variáveis que impactam a fruição dos indivíduos.

Outro conceito relevante para a arquitetura do DA é a criação de gatilhos, que surgem das Linhas de Design do projeto. Um gatilho é qualquer elemento adicionado ao longo da interação com o intuito de estimular ações dos indivíduos. Podem vir como elementos sonoros, visuais, de interface, interação ou mesmo chamadas de ação no próprio conteúdo. O DA considera basicamente dois tipos de gatilhos: Ação e Inércia. O primeiro estimula a percepção de valor ou utilidade em relação ao sistema audiovisual com o intuito de gerar maior engajamento, como por exemplo, sair do nível de Audiência ao assistir um programa de TV para compartilhar impressões online e se tornar um Sintetizador. Já o Gatilho de Inércia estimula o indivíduo à recepção do conteúdo por meio da atenção, análise ou interpretação. Um exemplo desse gatilho é quando o apresentador de um programa de TV convida a audiência a visualizar o resultado de sua colaboração online sobre algum aspecto do conteúdo. O DA será utilizado como base de descrição e criação deste trabalho à medida que propõe um conjunto de variáveis ou níveis de relações que devem ser consideradas por produtores de conteúdo dentro de um contexto de uso de mídia complexo. Nas próximas seções detalha-se a aplicação dos conceitos a partir do projeto de interação do sistema audiovisual e multi-tela Caatinga Viva.

4 MÉTODOS

A concepção do sistema multi-tela Caatinga Viva foi realizada essencialmente a partir dos elementos teórico-metodológicos do Design Audiovisual. A produção integra técnicas e processos comuns à criação audiovisual, engenharia de software, IHC e pedagogia. Neste trabalho serão descritos os processos relativos ao escopo da arquitetura de interação do sistema. De modo geral todo a construção pode ser dividida em pré, produção e pós.

A pré-produção consistiu essencialmente nas atividades de concepção do conteúdo e levantamento de requisitos para fins educacionais e de interação. Com o objetivo de otimizar a produção do experimento foi delimitada a obra televisiva Caatinga Viva como base do conteúdo em vídeo. O Caatinga Viva é uma obra audiovisual pré-existente a este artigo, de propriedade técnica da ONG Para'íwa. Dentre os sete programas produzidos pela ONG apenas o trecho relativo aos aspectos ambientais do Bioma Caatinga foi selecionado. Feito o primeiro filtro de conteúdo para a mídia primária, foi realizada uma visita técnica ao Centro Estadual Experimental de Ensino e Aprendizagem Sesquicentenário, localizado na cidade de João Pessoa - PB. Na ocasião, 14 professores da instituição realizaram voluntariamente uma análise do conteúdo audiovisual que definiu o requisitos educacionais do sistema. Como resultado deste processo, os professores colaboradores identificaram os conceitos de maior relevância presente no vídeo e fizeram uma lista de correlações para alimentar o dispositivo secundário.

²Disponível em: <<https://www.airconsole.com>>

Na etapa de produção foram realizados ajustes no material de vídeo, construção das estratégias de engajamento, interação multi-tela e a interface gráfica do sistema. O projeto de interação contempla a aplicação dos itens da arquitetura do DA como elementos de apoio tanto às premissas educacionais quanto aos requisitos de video-colaboração propostos. O detalhamento destes itens é apresentado na descrição conceitual do sistema. Por fim, a etapa de pós-produção contemplou o desenvolvimento da aplicação. O sistema foi implementado em linguagens web (HTML 5, CSS3 e Javascript) a partir de um micro-framework que configura o ambiente sobre um TCP de canais full-duplex entre browser e servidor, isto é, uma conexão websocket. Uma vez conhecidas as etapas de produção, vamos aprofundar os aspectos de conteúdo, interação e módulos do sistema.

5 O SISTEMA CAATINGA VIVA

A criação de um sistema audiovisual a partir da arquitetura do DA tem como etapa inicial a composição das Linhas de Design: Conteúdo, Experiência, Identidade e Motivação.

Linha de Conteúdo: a construção do conteúdo parte dos temas já existente no vídeo (mídia primária) e desdobra-se em abordagens complementares em uma aplicação multi-tela (mídia secundária).

Linha da Experiência: em paralelo à exibição do vídeo em uma grande tela (Projetor ou TV) uma aplicação web oferta conteúdos complementares. (ver figura 1). Ao longo deste processo gatilhos visuais, sonoros e de conteúdo estimulam nos estudante a alternância entre os dispositivos e níveis de engajamento.



Figura 1: Mídia primária (TV/projetor) e secundária (tablet)

Linha da Identidade: a identidade é retratada no Caatinga Viva essencialmente pelo contexto regional do tema, em que tanto a exposição didática no vídeo quanto os elementos visuais e avatares da aplicação secundária reforçam essa premissa.

Linha da Motivação: micro-interações com infográficos, mini jogos somados a estratégias de customização e personalização na aplicação web atuam na criação de valor e utilidade sobre o sistema audiovisual.

Integrado as demais linhas foram elaboradas micro-interações com os conteúdos que atuam como gatilhos de ação para outros níveis de engajamento.

5.1 Micro Interação: forma e conteúdo

A primeira lógica de micro interação é a exposição de conteúdos no formato de infografia. Nesta proposição estimula-se nos estudantes a ação de explorar os elementos de interface com o intuito de despertar o interesse pelo tema. Inicialmente a fruição deste tipo

de conteúdo equivale ao nível da audiência, isto é, a recepção em nível passivo ou individual da informação. Ao explorar a interface o estudante que está atualmente como audiência, pode alcançar o nível de sintetizador a partir do compartilhamento do conteúdo por mídias sociais ou mesmo comparando seus pontos com os demais colegas de classe dentro do ranking no próprio sistema.

A segunda lógica de micro interação para a aplicação dos dispositivos móveis é a oferta de conteúdos que desafiem o estudante a aplicar o tema que está sendo exibido na mídia primária (TV/Projetor). Essa abordagem estimula o estudante a sair de sua relação cognitiva de audiência para pôr a prova conceitos e acumular pontos dentro do próprio sistema. Ao interagir desta forma o estudante incorpora o estágio de player a sua fruição, isto é, amplia sua relação com o conteúdo a partir do exercício de novas competências e habilidades. Isso quer dizer que, após interagir (e conhecer a informação) ele criará um nível de identificação que permitirá o relacionamento mais avançado com o conteúdo, o que pode amparar tal aprendizado.



Figura 2: Mini jogo sobre intemperismo.

5.2 Macro interação: a fruição multi-tela

A construção de um ambiente multi-tela necessita da criação de estratégias de conteúdo e interação que possam garantir a fruição de mídia compartilhada, ao invés de uma sobreposição de dispositivos. A partir da arquitetura do Design Audiovisual (DA) foi possível planejar o sistema multi-tela Caatinga viva (CV) conforme Figura 3. Devido ao foco educacional do experimento, é natural que sua linha principal seja a do Conteúdo. Somado a isso, as Linhas da Experiência, Identidade e Motivação colaboram com a definição de conteúdo e gatilhos que promovem a alternância de dispositivos e níveis de engajamento.

O sistema multi-tela Caatinga Viva é composto por um vídeo com cinco minutos de duração e uma aplicação web. A aplicação é executada simultaneamente ao vídeo e fica disponível apenas enquanto este é reproduzido. Todo o conteúdo audiovisual foi separado em três blocos centrais (Pangeia, Intemperismo e Fosséis) de tal modo a modularizar os trechos de vídeo que correspondem exatamente aos conteúdos pedagógicos. Os itens essenciais a Linha de Conteúdo são destacados pela cor vermelha.

Motivados pela expansão da Experiência (itens amarelos), Identidade (itens azuis) e do próprio Conteúdo (itens vermelhos) surgem os gatilhos de ação do sistema audiovisual Caatinga Viva. Os gatilhos existem para motivar cada estudante a sair do vídeo e ir ao tablet, ou mesmo, compartilhar o resultado de sua interação através de mídias sociais. Além do indicador visual, a estratégia

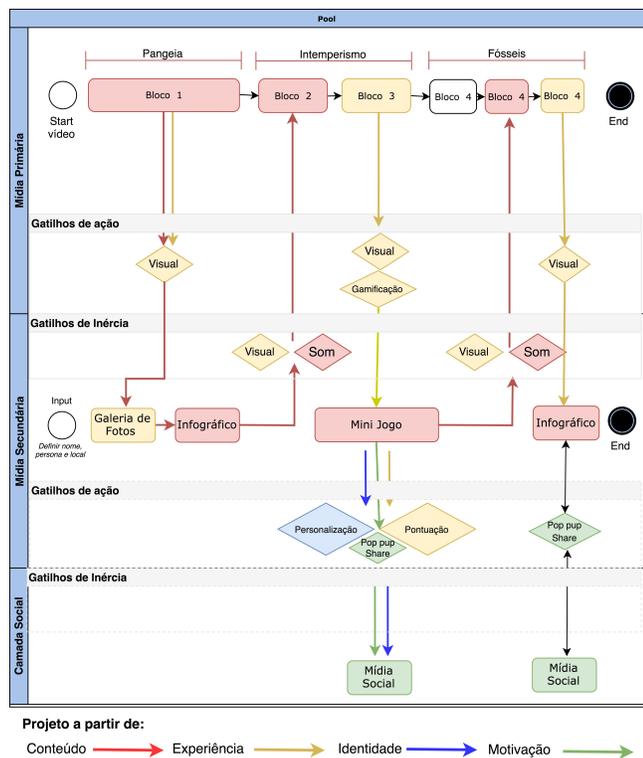


Figura 3: Arquitetura de conteúdo e interação

de Gamificação (pontuar escolhas) e a personalização de respostas foram utilizadas como gatilhos de ação. Já os gatilhos de inércia são apresentados tanto de forma visual quanto sonora, com o intuito de trazer a atenção do estudante de volta para um determinado momento em que o conteúdo é considerado indispensável pelo autor do sistema.

A arquitetura de conteúdo e interação apresentada na Figura 3 revela de forma visual como as Linhas de Design adotadas para o projeto repercutem na criação do conteúdo, funcionalidades do sistema e, principalmente, nos gatilhos que conduzem a alternância de dispositivo e nível de engajamento do indivíduo. Diante desta descrição percebe-se que a arquitetura do DA é um instrumento teórico-metodológico capaz de atender a produção de um ambiente multi-tela, uma vez que possibilita o planejamento da atenção e interação da fruição audiovisual a partir de itens adaptativos ao contexto de uso do sistema.

6 LIÇÕES APRENDIDAS

Conforme descrito na seção 2 existe uma escassez na literatura sobre estratégias que integrem aspectos de mídia e interação com foco na fruição dos indivíduos. Neste sentido, este trabalho se propôs experimentar a arquitetura do DA como base para a construção de um sistema multi-tela. Após a construção do experimento, experiências informais foram realizadas com professores e representantes da ONG. A pesquisa sobre a construção de sistemas audiovisuais multi-tela ainda encontra-se em andamento, com a perspectiva de

realização de testes com usuário para validação da usabilidade, escopo pedagógico e engajamento. Todavia, até o presente momento foram aprendidas as seguintes lições:

- As quatro linhas de Design permitem ao produtor e desenvolvedor estruturar com facilidade os requisitos de básicos conteúdo e interfaces interação de um sistema de videocolaboração multi-tela;
- O uso de sistemas para fins educacionais devem integrar metodologias específicas que complementem a definição de requisitos da Linha de Conteúdo;
- Os gatilhos apresentam-se como elemento funcional para projetar inércia e ação perante os níveis de engajamento e dispositivos ofertados;
- O sistema deve ofertar mais de um formato de gatilho (seja para inércia ou ação). Uma vez que um ponto de interesse para um grupo de pessoas não é necessariamente o mesmo para outro grupo;

Os impactos destes elementos ficam a cargo de estudos futuros, na continuação desta pesquisa.

AGRADECIMENTOS

Agradecemos ao professor Cléber de Moraes (UFPP) e a desenvolvedora Camila Cavalcante pelo apoio técnico e científico.

REFERÊNCIAS

- [1] Valdecir Becker, Daniel Gambaro, and Thais Saraiva Ramos. 2017. Audiovisual design and the convergence between HCI and audience studies. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 10271. 3–22. https://doi.org/10.1007/978-3-319-58071-5_1
- [2] Bruce Block. 2010. *A narrativa visual: criando a estrutura visual para cinema, TV e mídias digitais*. Elsevier, São Paulo. 312 pages.
- [3] Nathan Nascimento Cirino. 2012. *CINEMA INTERATIVO : Problematizações de linguagem e roteirização*. Ph.D. Dissertation. Universidade Federal de Pernambuco, Recife.
- [4] Francisco Machado Filho and Leire Mara Bevilacqua. 2014. Pré-requisitos para produção de conteúdo em segunda tela para televisão aberta. In *XXXVII Congresso Brasileiro de Ciências da Comunicação*. Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, Foz do Iguaçu, PR, 1–15.
- [5] Vicente Gosciola. 2008. *Roteiro para as novas mídias - Do cinema às mídias interativas* (2 ed.). Senac - SP, São Paulo. 277 pages.
- [6] Henry Jenkins, Sam Ford, and Joshua Green. 2014. *Cultura da conexão: Criando valor e significado por meio da mídia propagável*. Aleph, São Paulo.
- [7] Vanessa Regina Margareth Lima Maíke, Laurindo de Sousa, Britto Neto, Simone Klein Goldenstein, and Maria Cecília Calani Baranauskas. 2015. *Heuristics for NUI Revisited and Put into Practice*. Springer International Publishing, Cham. 317–328 pages. <https://doi.org/10.1007/978-3-319-20916-6>
- [8] Britta Meixner. 2017. Hypervideos and Interactive Multimedia Presentations. *Comput. Surveys* 50, 1 (2017), 1–34. <https://doi.org/10.1145/3038925>
- [9] G Millerson and J Owens. 2008. *Video production handbook - fourth edition*. Vol. 4. Focal Press, Burlington, USA. 9823 pages. [https://doi.org/10.1002/1521-3773\(20010316\)40:6<9823::AID-ANIE9823>3.3.CO;2-C](https://doi.org/10.1002/1521-3773(20010316)40:6<9823::AID-ANIE9823>3.3.CO;2-C)
- [10] Luís Nogueira. 2008. *Narrativas filmicas e videojogos*. Vol. 1. LabCom, Beira interior. 388 pages. <https://doi.org/10.1017/CBO9781107415324.004> arXiv:arXiv:1011.1669v3
- [11] Luís Nogueira. 2015. Cinema múltiplo, figuras, temas, estilos, dispositivos. (2015), 185 pages.
- [12] Nitin Sawhney, David Balcom, and Ian Smith. 1997. Authoring and navigating video in space and time. *IEEE Multimedia* 4, 4 (1997), 30–39. <https://doi.org/10.1109/93.641877>
- [13] Jorge Pedro Sousa. 2006. *Elementos de Teoria e Pesquisa da Comunicação e dos Mídia* (2 ed.). Edições Universidade Fernando Pessoa, Porto, Portugal (PT).
- [14] Rafael Toscano, Valdecir Becker, Laiz Ferreira, Samara Coutinho, and Leonardo Burgos. 2017. Arquitetura de design colaborativo para imersão temporal e espacial em vídeos de altíssimas resoluções e HFR. In *O futuro da videocolaboração: perspectivas*. Sociedade Brasileira de Computação (SBC), Porto Alegre, RS, Chapter 1, 13–53. <http://www.inf.ufprgs.br/webmedia2017/wp-content/livro>

VideoRecognition - Uma proposta de serviço para reconhecimento de elementos de vídeo em larga escala

Antonio J. G. Busson
TeleMídia/PUC-Rio
busson@telemidia.puc-rio.br

Álan L. V. Guedes
TeleMídia/PUC-Rio
alan@telemidia.puc-rio.br

Gabriel N. P. dos Santos
Faculdade ISL | Wyden
gabrieainha@gmail.com

Carlos de Salles Soares Neto
TeleMídia/UFMA
csalles@deinf.ufma.br

Ruy Luiz Milidiú
LEARN/PUC-Rio
ruy@inf.puc-rio.br

Sergio Colcher
TeleMídia/PUC-Rio
colcher@inf.puc-rio.br

ABSTRACT

Deep Learning research has allowed significant advancement of various segments of multimedia, especially in tasks related to speech processing, hearing and computational vision. However, some video services are still focused only on the traditional use of media (capture, storage, transmission and presentation). In this paper, we discuss our ongoing research towards a DLaaS, *i.e.* Deep Learning as a Service. This way, we present the state of art in video classification and recognition. Then we propose the *VideoRecognition* as DLaaS to support the tasks such as: image classification and video scenes, object detection and facial recognition. We discuss the usage of the proposed service in the context of the video@RNP repository. Our main contributions consist on discussions over the state of art and its usage in nowadays multimedia services.

KEYWORDS

Video Recognition, Deep Learning, CNN, Video repositories

1 INTRODUÇÃO

Nos últimos anos, o *Deep Learning* (DL) permitiu significativo avanço de vários segmentos da multimídia, principalmente em tarefas relacionadas a processamento de fala, audição e visão computacional [17]. Plataformas como IBM Watson¹ e Microsoft Azure ML² já oferecem DLaaS (*Deep Learning as a Service*), permitindo que sistemas multimídia (e.g. Helpicto [5], PersonalizedTV [15], ShrewsburyMuseum [11]) possam incorporar novas funcionalidades baseadas em aprendizagem e reconhecimento de padrões, o que os leva para a categoria de IIMS (*Intelligent Interactive Multimedia Systems*).

Na contramão, muitos serviços multimídia usados no Brasil (e.g. Video@RNP³, Videoaula@RNP⁴) ainda estão focados apenas no uso tradicional das mídias (captura, armazenamento, transmissão e apresentação). O uso de um serviço de reconhecimento está fortemente ligado a melhora na busca e indexação do conteúdo através da

¹<https://www.ibm.com/watson/>

²<https://azure.microsoft.com/pt-br/services/machine-learning-studio/>

³<http://video.rnp.br/portal/home>

⁴<http://www.videoaula.rnp.br/portal/home>

In: V Workshop "O Futuro da Videocolaboração" (WCT-Video 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pósteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.

© 2018 SBC – Sociedade Brasileira de Computação.

ISBN 978-85-7669-435-9.

descoberta de metadados mais precisos, mas também pode ser útil para a criação de funcionalidades sofisticadas, como por exemplo: enriquecimento de conteúdo, reconhecimento facial, monitoração de *streaming*, detecção de conteúdo impróprio, etc. Nesse sentido, é interessante que exista um serviço dedicado à aprendizagem e reconhecimento de conteúdo multimídia, que tenha baixo custo e que mantenha o sigilo dos dados.

Neste artigo, propomos o *VideoRecognition*, um serviço para reconhecimento de elementos de vídeo em larga escala. O *VideoRecognition* faz uso de tecnologias abertas de *Deep Learning* para suportar tarefas de: classificação de imagem e cenas de vídeo, detecção de objetos e reconhecimento facial. A arquitetura do serviço foi planejada para cumprir três requisitos básicos: (1) permitir que analistas possam especificar arquiteturas de DL e padrões de dados para tarefas de reconhecimento específicas; (2) oferecer uma ferramenta para que usuários possam anotar conteúdo multimídia e gerar *datasets*; (3) oferecer uma interface para integração com outros serviços multimídia. Dessa forma, possibilitando que outros serviços possam incorporar funcionalidades baseadas em reconhecimento automático.

O restante desse artigo está organizado como segue. A Seção 2 descreve o *estado-da-arte* nas tarefas de classificação de imagem e vídeo, detecção de objetos e reconhecimento facial. Em seguida, a Seção 3 apresenta a proposta deste artigo. A Seção 4 apresenta um cenário de uso para integração da proposta com o serviço Video@RNP. Por fim, a Seção 5 contém as considerações finais.

2 ESTADO DA ARTE

Métodos baseados em *Deep Learning* se tornaram o estado da arte em vários segmentos da área de sistemas multimídia. Nesta seção, são apresentados os modelos baseados em *Deep Learning* que já são considerados maduros suficientes para o desenvolvimento de serviços de reconhecimento de elementos de vídeo. Na Subseção 2.1 é apresentado o *estado-da-arte* para classificação de imagem e vídeo. Em seguida, na Subseção 2.2 é apresentado o *estado-da-arte* para as tarefas de detecção de objetos e reconhecimento facial.

2.1 Classificação de imagem e vídeo

Na multimídia a tarefa de classificação consiste em mapear um conteúdo de mídia em uma ou mais categorias distintas. Arquiteturas de *Deep Learning* baseadas em CNNs (*Convolutional neural network*) ou ConvNets se tornaram o principal método usado para reconhecimento de padrões áudio-visuais. Tipicamente o treinamento de

CNNs é feito de maneira supervisionada, e são treinadas em *datasets* que contém milhares/milhões de mídias e classes relacionadas. Durante o treinamento, as CNNs aprendem a hierarquia de *features* que são aplicadas a mídia de entrada para que seja possível realizar a classificação do seu conteúdo.

Desde a vitória da CNN AlexNet [12] no desafio ImageNet 2012 [20], surgiram novas arquiteturas baseadas em CNN cada vez mais precisas. A vencedora do ImageNet 2014, por exemplo, foi a CNN InceptionNet [23], que propôs o uso do bloco Inception, um bloco que usa vários filtros de diferentes tamanhos no mesmo nível para resolver o problema de localização da informação em imagens. No ano seguinte, a rede ResNet [7] foi a vencedora do ImageNet 2015, e introduziu o conceito de conexões residuais, que aumentou a performance e reduziu o tempo de treinamento das CNNs. Mais tarde foi desenvolvida a arquitetura Inception-Resnet [22], que combina os blocos Inception com as conexões residuais. Essa arquitetura é popular e é a base para muitas outras arquiteturas de CNN para extração de *features*.

A arquitetura SE-Net (*Squeeze-and-Excitation Network*) [10] é o *estado-da-arte* na tarefa de classificação de imagens, obtendo 2.25% de erro top-5 no ImageNet 2017. Assim como o InceptionNet, a SE-Net propõe um novo tipo de bloco chamado SE, que melhora o poder de representação da rede ao destacar as interdependências entre os canais da imagem e seus mapas de *features*. Para isso, a SE-Net usa um mecanismo que permite que a rede possa fazer uma recalibração de *features*, através do qual, usa informações globais para enfatizar as *features* mais informativas e suprimir as *features* menos úteis.

Diferente de imagens, vídeos não possuem apenas informação visual, mas também auditiva. Métodos atuais para classificação de vídeo geralmente representam os vídeos pelas *features* audiovisuais extraídas dos seus quadros, seguido pela agregação dessas *features* sobre o tempo. Métodos usados para extração de *features* incluem modelos de CNN pré-treinados em *datasets* de larga escala. O YouTube8M [1], por exemplo, usa a rede Inception-Resnet pré-treinada no ImageNet para extrair *features* visuais, e usa a rede Audio-VGG [8] pré-treinado no Audioset [6] para extrair *features* auditivas. Já métodos para agregação de *features* incluem métodos sofisticados de *pooling* como na CNN NetVLAD [2] ou usando modelos recorrentes baseados em GRU [3] ou LSTM [9]. A arquitetura Gated-NetVLAD [16] conseguiu obter 84% de GAP no desafio YouTube8M 2017⁵, que contém mais de 6 milhões de vídeos distribuídos em 3862 classes de etiquetas.

2.2 Detecção de objetos e reconhecimento facial

A tarefa de detecção de objetos consiste em localizar e classificar objetos que estão dentro da mídia. A arquitetura YOLO (*You Only Look Once*) [18] é considerada o *estado-da-arte* na tarefa de detecção de objetos. Sua última versão, chamada YOLOv3 [19] obteve um mAP de 57.9% no dataset COCO (*Common Objects in Context*) [14]. O YOLO é ideal para aplicações de vídeo em tempo-real, visto que é o modelo de detecção de objetos baseado em CNN mais rápido da literatura, chegando a rodar próximo de 30 FPS na GPU Pascal Titan

⁵<https://research.google.com/youtube8m/>

X⁶. O YOLO divide a imagem em várias regiões e prediz uma caixa delimitadora de objeto para cada região. As coordenadas da caixa delimitadora podem ser usadas para destacar ou segmentar o objeto da imagem. As imagens segmentadas, por sua vez, podem ser usadas em outros tipos de tarefa, como por exemplo, de reconhecimento facial.

A tarefa de reconhecimento facial tenta responder a pergunta “Quem é esta pessoa?”. Tipicamente, a identificação de imagens de face é feita pela comparação da similaridade entre uma ou mais imagens de indivíduos já conhecidos. A arquitetura FaceNet [21] é o *estado-da-arte* para a tarefa de reconhecimento facial. Esta arquitetura tem uma acurácia de 99.6% no *dataset* LFW (*Labeled Faces in the Wild*) [13]. Dada duas imagens de face, o FaceNet extrai as *features* faciais a partir da ativação linear da última camada densa da rede Inception-Resnet. Em seguida, a similaridade entre as imagens de face é calculada pela distância euclidiana entre os vetores de *features*, se a distância for menor que um limiar, então presume-se que as imagens de face sejam da mesma pessoa. Ao usar o FaceNet em conjunto com o YOLO é possível realizar reconhecimento facial em vídeo e *live streaming*.

3 PROPOSTA DE TRABALHO

Nesta seção é apresentado o *VideoRecognition*, um serviço que visa oferecer interfaces para anotação de conteúdo e reconhecimento de elementos audiovisuais. A ideia é permitir que outros serviços multimídia possam utilizar essas interfaces para incorporar funcionalidades mais sofisticadas a suas próprias arquiteturas. Inicialmente o *VideoRecognition* é restrito a três tipos de tarefas: classificação de imagem e cenas de vídeo, detecção de objetos e reconhecimento facial. Adiante é apresentada a arquitetura do *VideoRecognition*, e em seguida, uma discussão de como outros serviços multimídia podem se beneficiar das funcionalidades do *VideoRecognition*.

A arquitetura do *VideoRecognition* é ilustrada na Figura 1. O grupo *Client Services* contém serviços que utilizam o *VideoRecognition*, como exemplos citamos: Um repositório de Vídeo (*RV*), um serviço de conferência de vídeo (*VideoConference*), e por último, um editor de vídeo (*VideoEditor*). O *VideoRecognition* possui duas interfaces de acesso: *annotationIF* e *recognitionIF*. A interface *annotationIF* permite que os serviços possam alimentar o *VideoRecognition* com mídias anotadas, em outras palavras, permite a criação de *datasets*. Posteriormente os *datasets* são utilizados para o treinamento dos modelos de reconhecimento. Por sua vez, a interface *recognitionIF* permite que os serviços possam utilizar os modelos de reconhecimento já treinados.

O *VideoRecognition* contém uma pilha de modelos, onde cada um é especializado em um tipo de tarefa e dedicado a um serviço externo. Esses modelos são projetados por analistas que especificam a arquitetura de DL (dentre as descritas na Seção 2), as classes, tipo de tarefa e qual serviço pode utilizá-lo. Ao acessar uma interface, o serviço cliente deve especificar qual modelo da pilha deve ser utilizado. Cada modelo da pilha possui dois sub-módulos: *train* e *recognize*. O módulo *train* usa as informações do *dataset* para realizar o treinamento da arquitetura, por sua vez o módulo *recognize* utiliza os pesos do treinamento para realizar as predições solicitadas através da interface *recognitionIF*.

⁶<https://www.nvidia.com/pt-br/geforce/products/10series/titan-x-pascal/>

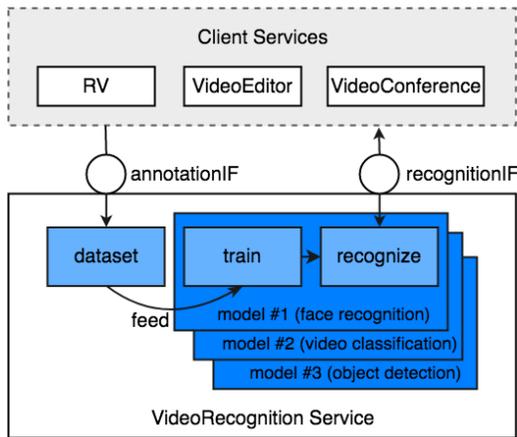


Figura 1: Arquitetura do VideoRecognition.

Ao utilizarem o VideoRecognition, os Client Services podem prover, dentre outras, as seguintes funcionalidades (F):

- **F.1. Busca utilizando mídia de referência.** Essa funcionalidade visa permitir usuários de um Client Service proativamente encontrar conteúdos de mídias que pertencem a uma mesma classificação. Nela, o usuário utiliza um conteúdo de mídia como parâmetro de busca. Em um RV, por exemplo, o usuário pode inserir a foto da Grécia antiga e, em seguida, o VideoRecognition pode listar vídeo aulas sobre a Grécia antiga. Adicionalmente, um criador de vídeo utilizando um VideoEditor pode visualizar quais outros conteúdos presentes em um RV são relacionados ao vídeo que está editando. Isso permite enriquecer o seu conteúdo de vídeo, seja com novos links, ou novos conteúdos de mídias.
- **F.2. Reconhecimento facial.** Essa funcionalidade visa permitir usuários de um Client Service proativamente encontrarem conteúdos de mídias que possuem a mesma pessoa. Nela, o usuário utiliza um conteúdo de mídia com pessoa específica como parâmetro de busca. Em um RV, por exemplo, o usuário pode inserir a foto de um professor específico para o RV listar vídeos onde aparece o professor desejado.
- **F.3. Recomendação.** Essa funcionalidade está relacionada com a capacidade de um Client Service automaticamente listar conteúdos de mesma classificação ou com uma mesma pessoa para seus usuários. Em uma RV, por exemplo, a visualização de um vídeo pode ser acompanhada com a lista de vídeos relacionados.
- **F.4. Detectar conteúdo impróprio.** Essa funcionalidade visa prevenir a presença de conteúdos impróprios em Client Services. Essa funcionalidade requisita um treinamento prévio de classificação de conteúdos que são considerados impróprios (e.g., cenas de sexo, uso de drogas, gore). Um Client Service pode solicitar que o VideoRecognition analise o conteúdo da mídia para verificar se o conteúdo é impróprio.
- **F.5. Busca por trecho de interesse.** Essa funcionalidade é relacionada com a busca ou segmentação temporal de trechos de vídeo segundo um critério de filtragem estabelecido

pelo usuário de um Client Service. Em um VideoEditor, por exemplo, o usuário pode solicitar que o VideoRecognition segmente todos os trechos do vídeo onde aparecem paisagens. Em um cenário de RV, o usuário pode fazer buscas por trechos de vídeo específicos, onde o VideoRecognition pode, por exemplo, listar apenas cenas de vídeo onde aparecem carros esportivos.

- **F.6. Sincronismo com eventos de reconhecimento.** Essa funcionalidade está relacionada com a capacidade de sincronizar elementos de mídia com eventos de reconhecimento. Nesse caso, o usuário de um Client Service pode estabelecer regras de reconhecimento para iniciar ou terminar a apresentação de uma ou mais mídias. Por exemplo, durante a transmissão de uma palestra pelo VideoConference, uma janela com fotos e texto biográfico sobre o palestrante podem estar disponíveis enquanto o palestrante estiver em cena. o VideoRecognition pode notificar o serviço VideoConference sobre presença e identidade do palestrante para que o conteúdo multimídia correto seja apresentado.

4 CASO DE USO NO VIDEO@RNP

Com o propósito de avaliar a proposta, nesta seção é apresentado um cenário de uso para integração do VideoRecognition com um serviço existente. Mais precisamente, propomos a integração com o RV vídeo@RNP. Segundo Ciuffo *et. al.* [4], o vídeo@RNP é um serviço para disponibilização e armazenamento de conteúdo audiovisual de instituições acadêmicas que tem como diferencial sua infraestrutura de rede, um tipo de CDN (Content Delivery Network) restrita ao tráfego audiovisual. Em particular, inicialmente é proposto que esse serviço seja estendido com as funcionalidades F.2 e F.3.

Relacionado a F.2, identifica-se que uma das limitações desse serviço é indexação de vídeos, e consequentemente a sua busca, de conteúdo. Por exemplo, o vídeo⁷ do professor Marcos Bagno sobre gramática realizado na Faculdade de Ciências e Letras (FCL) da UNESP possui as keywords: "Marcos", "Bagano", "FCL" e "gramática". Percebemos que o professor não está corretamente indexado. Acreditamos que essas keywords tenham sido retiradas automaticamente da descrição do vídeo. Essa limitação poderia ser contornada com o uso do VideoRecognition.

O vídeo@RNP poderia acessar o módulo de recognition para realizar a indexação de pessoas nos vídeos. Entretanto, o recognition precisa ser treinado para identificar pessoas específicas (e.g. quem é Marcos Bagno). Para isso, primeiro o vídeo@RNP necessita solicitar ao módulo recognition para segmentar faces. Em seguida, através de uma interface de anotação, um usuário de administração do vídeo@RNP poderia anotar faces reconhecidas e enviar para o módulo annotation. A Figura 2 ilustra uma mockup dessa interface. Ela indica o vídeo analisado e a segmentação de pessoas durante a apresentação do vídeo (Figura 2-A). Adicionalmente, ela apresenta uma lista de reconhecimentos de faces (Figura 2-B). Tais reconhecidos podem indicar novas faces não conhecidas (Figura 2-C) ou faces já conhecidas que necessitam de confirmação (Figura 2-D).

Relacionado a F.3., identifica-se que o vídeo@RNP realiza apenas uma busca facetada, ou seja, usuários apenas navegam em conjuntos de vídeos definidos pelas keywords inseridas na criação dos vídeos.

⁷<http://video.rnp.br/portal/video.action?itemId=26242>

Esse serviço poderia se utilizar do módulo recongnition para gerar uma lista de vídeos relacionados a semântica do conteúdo do vídeo visualizado. O critério semântica poderia pode a classificação do vídeo (e.g. aulas de gramáticas) ou a presença de uma pessoa específica (e.g. professor Marcos Bagno).

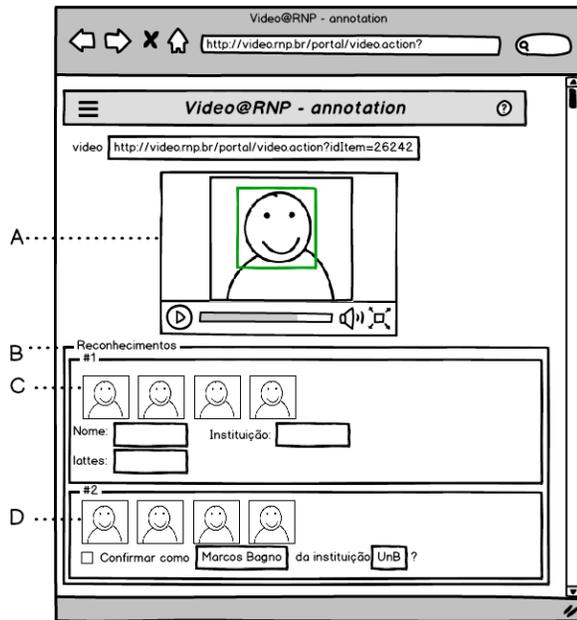


Figura 2: Esboço da interce de anotação do video@RNP com suporte do VideoRecognition

5 CONSIDERAÇÕES FINAIS

Este trabalho apresenta uma visão geral do estado da arte de atividades de reconhecimento de elementos audiovisuais em vídeos. Com o objetivo de preencher a lacuna desses atividades em serviços certos serviços multimídia, apresentamos o VideoRecognition como sendo uma proposta de DLaaS (Deep Learning as a Service). Ele visa oferecer suporte a outros serviços de vídeo como RV, editor de vídeo e vídeo conferência. Nesse sentido, apresentamos uma visão arquitetural do VideoRecognition e discutimos para estender outros serviços de vídeo. Em particular, discutimos um caso de uso sobre RV chamado video@RNP e apresentamos um mockup da ferramenta de anotação. Nossas principais contribuições consistem em discussões sobre o estado da arte e no uso do VideoRecognition em serviços multimídia atuais.

O serviço faz parte de uma pesquisa em andamento. Citamos como trabalhos futuros: (a) explorar novos métodos de Deep Learning dedicados a outras modalidades de mídia que ainda não são suportadas pelo VideoRecognition, como áudio e legendas de vídeo; (b) enriquecer o VideoRecognition com novos tipos de tarefas, como por exemplo, detecção de atividades em vídeo; (c) promover sessões de design participativo com desenvolvedores e administradores de outros serviços multimídia a fim de identificar e validar requisitos para o VideoRecognition.

REFERÊNCIAS

- [1] Sami Abu-El-Hajja, Nisarg Kothari, Joonseok Lee, Apostol (Paul) Natsev, George Toderici, Balakrishnan Varadarajan, and Sudheendra Vijayanarasimhan. 2016. YouTube-8M: A Large-Scale Video Classification Benchmark. In *arXiv:1609.08675*. <https://arxiv.org/pdf/1609.08675v1.pdf>
- [2] Relja Arandjelovic, Petr Gronat, Akihiko Torii, Tomas Pajdla, and Josef Sivic. 2016. NetVLAD: CNN architecture for weakly supervised place recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 5297–5307.
- [3] Kyunghyun Cho, Bart Van Merriënboer, Dzmitry Bahdanau, and Yoshua Bengio. 2014. On the properties of neural machine translation: Encoder-decoder approaches. *arXiv preprint arXiv:1409.1259* (2014).
- [4] Leandro Ciuffo, Marcelino Cunha, Christian Miziara de Andrade, Graciela Leopoldo Martins, Jean Carlo Faustino, Rafael Valle, Fausto Vetter, Antônio Carlos Nunes, and Helder Vitorino. Mapeamento de soluções de video colaboração da RNP. In *Anais do XXIII Simpósio Brasileiro de Sistemas Multimídia e Web: Workshop do CT-Vídeo (Comitê Técnico de Prospecção Tecnológica em Videocolaboração)*. 337–368.
- [5] Equadex. 2018. Helpicto. (2018). <http://www.helpicto.com/> Accessed: 2018-05-18.
- [6] Jort F Gemmeke, Daniel PW Ellis, Dylan Freedman, Aren Jansen, Wade Lawrence, R Channing Moore, Manoj Plakal, and Marvin Ritter. 2017. Audio set: An ontology and human-labeled dataset for audio events. In *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*. IEEE, 776–780.
- [7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.
- [8] Shawn Hershey, Sourish Chaudhuri, Daniel P. W. Ellis, Jort F. Gemmeke, Aren Jansen, Channing Moore, Manoj Plakal, Devin Platt, Rif A. Saurous, Bryan Seybold, Malcolm Slaney, Ron Weiss, and Kevin Wilson. 2017. CNN Architectures for Large-Scale Audio Classification. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. <https://arxiv.org/abs/1609.09430>
- [9] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.
- [10] Jie Hu, Li Shen, and Gang Sun. 2017. Squeeze-and-excitation networks. *arXiv preprint arXiv:1709.01507* 7 (2017).
- [11] Martin Kearn and Martin Beeby. 2017. Using Cognitive Services to make museum exhibits more compelling and track user behavior. (2017). <https://microsoft.github.io/techcasesstudies/cognitive%20services/2017/0/8/04/BlackRadley.html> Accessed: 2018-05-18.
- [12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2012. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*. 1097–1105.
- [13] Gary B. Huang Erik Learned-Miller. 2014. *Labeled Faces in the Wild: Updates and New Reporting Procedures*. Technical Report UM-CS-2014-003. University of Massachusetts, Amherst.
- [14] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. 2014. Microsoft coco: Common objects in context. In *European conference on computer vision*. Springer, 740–755.
- [15] Julio Wong Mathews Thomas, Janki Vora and Satish Sadagopan. 2016. Use cases for industry cognitive solutions. (2016). <https://www.ibm.com/developerworks/library/cc-cognitive-media-telco-2-trs/index.html> Accessed: 2018-05-18.
- [16] Antoine Miech, Ivan Laptev, and Josef Sivic. 2017. Learnable pooling with Context Gating for video classification. *arXiv preprint arXiv:1706.06905* (2017).
- [17] Kaoru Ota, Minh Son Dao, Vasileios Mezaris, and Francesco GB De Natale. 2017. Deep learning for mobile multimedia: A survey. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 13, 3s (2017), 34.
- [18] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. 2016. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 779–788.
- [19] Joseph Redmon and Ali Farhadi. 2018. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767* (2018).
- [20] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. 2015. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision* 115, 3 (2015), 211–252.
- [21] Florian Schroff, Dmitry Kalenichenko, and James Philbin. 2015. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 815–823.
- [22] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alexander A Alemi. 2017. Inception-v4, inception-resnet and the impact of residual connections on learning. In *AAAI*, Vol. 4. 12.
- [23] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. 2015. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 1–9.

VideoViz - Uma proposta de ferramenta de visualização de *logs* de interação em *players* de vídeo

André Luiz de B. Damasceno¹, Álan L. V. Guedes¹, Sergio Colcher¹, Simone D. J. Barbosa¹

Dilson J. L. Rabêlo², Elza B. Monier²

¹Pontifícia Universidade Católica do Rio de Janeiro – Rio de Janeiro, RJ – Brasil

²UNA-SUS/UFMA – São Luís, MA – Brasil

{andre,alan}@telemidia.puc-rio.br,{colcher,simone}@inf.puc-rio.br,{dilsonrabelo.unasus,bernardes.elza}@gmail.com

ABSTRACT

The usage of technologies for content authoring, storage, distribution and presentation supported the arise of Virtual Learning Environments (VLEs) such as Moodle and MOOCs (e.g., Udacity, Coursera, EdX). All students' interaction in these environments can be stored in logs. For instance, video is one of the resources has a greater focus on VLEs. In this paper, we present a state-of-the-art overview of interaction analysis on video players. We noticed that teachers would like to know how and when students interact with videos. In order to fill this gap, we propose VideoViz, a tool to analyse logs from interaction on video players.

KEYWORDS

Learning Analytics, Educational Data Mining, MOOC, AVA, e-learning, Video, engagement, behavior, interaction, visualization, tool

1 INTRODUÇÃO

Estudantes e professores atualmente utilizam tecnologias que permitem novas maneiras de aprender e ensinar. Apesar da educação a distância (EaD) não ser novidade, havendo registros do século XIX de cursos por correspondência [15], foi a partir da utilização de computadores e da evolução da Internet que permitiu ao aluno participar de forma mais ativa e gerou novos paradigmas nessa modalidade de ensino. Um dos motivos para o crescimento da oferta e demanda de cursos a distância é a conveniência dos alunos determinarem seu ritmo de aprendizagem e a possibilidade deles participarem de cursos sem serem limitados pela distância geográfica. Em particular, o uso de tecnologias de autoria, armazenamento, distribuição e apresentação de conteúdo contribuíram para o surgimento de Ambientes Virtuais de Aprendizagem (AVAs) como MOOCs (*Massive Open Online Courses*) (e.g., Udacity¹, Coursera², EdX³).

AVAs como o Moodle⁴ não são restritos aos cursos inteiramente *online*. Conforme evidenciado por Damasceno et al. [6], há casos em que AVAs são utilizados em cursos presenciais. Nessa modalidade de ensino, também conhecida como *blended learning*, o professor pode complementar sua aula com os seguintes recursos: postagem de

notícias, *links* de materiais da Internet, ferramentas de comunicação (e.g., fórum, chat), *quiz* e *wiki*. Outra vantagem dos AVAs é a sua capacidade de coletarem e armazenarem toda interação dos estudantes em formato de *logs*. Através da análise desses *logs* é possível obter informações sobre padrões, comportamento, desempenho e características dos estudantes [7, 14].

Educational Data Mining (EDM) e *Learning Analytics* (LA) são áreas de pesquisa com objetivos e interesses similares na exploração de *dados* (e.g., *logs*) em EaD. Ambas utilizam detecção automática de informação e análise humana através de visualizações. Contudo, a EDM utiliza a análise humana como ferramenta para melhorar a detecção automática de informação, enquanto a LA busca potencializar a análise humana através de ferramentas que realizam a detecção automática de informação [17].

Em [6], os autores apresentam um mapeamento sistemático da literatura sobre trabalhos que analisam *logs* de interação de alunos em AVAs. Além disso, entrevistas conduzidas com professores mostram quais informações dos alunos eles consideram relevantes e quais são os requisitos necessários para ajudá-los na análise dessas informações. Os autores também observaram que a maioria dos artigos encontrados visam identificar padrões de comportamento e performance dos estudantes. Entre esses trabalhos, foram encontrados 20 ferramentas para auxiliar o professor na análise de *logs*. Porém, nenhuma delas preenche todos os requisitos mencionados pelos professores.

Os professores consideram importante extrair informação a partir da interação e *feedbacks* dos alunos nos sistemas de EaD [6]. Contudo, a maioria dos professores obtêm informações limitadas, como dados de último acesso dos estudantes e postagens dos alunos nos fóruns dos AVAs. Em geral, os professores não recebem treinamento para extrair informações dos *logs*. Por isso, eles sentem necessidade de ferramentas que analisam *logs* de acesso e interação dos estudantes, tanto com os materiais, quanto com os recursos dos AVAs como fórum, *player* de vídeo, *quiz* e *ebooks*.

Nesse contexto, este artigo apresenta nossa pesquisa em andamento sobre análise de *logs* de vídeos. Mais precisamente, apresentamos uma visão geral do estado da arte de análise dos *logs* de interação em *players* de vídeos. Em particular, essa visão geral também apresenta o que professores gostariam de saber, como quando e de que forma os alunos interagem com os vídeos. Com o objetivo de preencher a lacuna de ferramentas para auxílio na análise desses *logs*, apresentamos uma proposta de ferramenta chamada *VideoViz*.

O restante deste artigo está organizado como a seguir. A Seção 2 discute trabalhos que compartilham nossa motivação de análise de vídeos. A Seção 3 apresenta nossa proposta de ferramenta de

¹<https://udacity.com>

²<https://coursera.org>

³<https://edx.org>

⁴<https://moodle.org/>

análise de vídeos chamada VideoViz. Por fim, a Seção 4 discute considerações finais e trabalhos futuros.

2 ESTADO DA ARTE

Vídeo é um dos recursos que tem ganhado destaque nos cursos de EaD [16] e o *streaming* é a forma dominante de acesso a eles [10]. Segundo alguns professores, a videoaula é o formato de conteúdo que os alunos mais gostam de acessar, porque além de permitir que a explicação de um assunto possa ser vista mais de uma vez, em muitas vezes é o primeiro recurso utilizado para entender um determinado assunto [6]. Portanto, pesquisas como autoria de videoaula e interações com *players* de vídeo têm tido relevância na área de Informática na Educação (IE). Nesta seção apresentamos o estado da arte das pesquisas em IE envolvendo vídeo, organizadas nas seguintes perspectivas: 2.1 análise de engajamento em vídeos; e 2.2 visualização de informação.

2.1 Análise de engajamento em vídeos

Em IE, engajamento é um termo associado a mais de um significado (e.g., número de materiais acessados, acessos aos AVAs, atividades realizadas, interações e postagens nos fóruns). Contudo, em pesquisas com vídeos engajamento é associado a retenção ou tempo de visualização do vídeo [1, 8]. Um dos principais problemas identificados pelos professores é o engajamento dos alunos nos vídeos, pois geralmente eles não assistem uma videoaula inteira [8, 9, 12]. A retenção dos alunos nos vídeos é um dado importante a ser levado em consideração, pois o tempo que os alunos passam assistindo uma videoaula está correlacionado com o desempenho obtido por eles nos cursos [13, 21].

Muitos professores possuem experiência no desenvolvimento de material educacional como vídeos, *slides* e áudios. Por isso, em muitos cursos eles acabam desempenhando também o papel de autores de conteúdo [6]. Segundo eles, há indícios de que a retenção dos alunos nos vídeos está relacionada com a autoria da videoaula. Existem pelo menos 13 estilos diferentes de apresentação dessas videoaulas [4], e algumas pesquisas relatam que o formato de *slides* não obtém um bom resultado [6, 12]. Corroborando com esses indícios, Guo et al. [8] verificaram que videoaulas no formato *Khan academy* e *Talking-head* são os que mais favorecem o engajamento dos alunos, filmagens de aulas presenciais têm um baixo índice de retenção dos alunos e as videoaulas que passam por um processo de pré-produção apresentam resultados melhores na permanência dos alunos. Além disso, os alunos se engajam mais com vídeos filmados informalmente com o professor sentado em sua mesa, quando comparados com vídeos produzidos no formato de produção de alto custo de TV [8].

Em entrevistas [6], alguns professores citaram como fatores que influenciam o aluno a assistir uma videoaula: (i) a proximidade da prova, (ii) a produção e a edição do vídeo, (iii) o formato de exibição do conteúdo e (iv) a maneira como o professor se apresenta e se expressa na videoaula, pois professores que falam devagar ocasionam um baixo engajamento nas videoaulas [8]. A duração do vídeo foi mais um fator mencionado por esses professores. Segundo eles, o aluno não costuma assistir vídeos longos e, quando acessa, não assiste o vídeo todo. Contudo, não houve consenso sobre a duração ideal de uma videoaula (variando de 5 a 30 minutos de

duração máxima). A variação do engajamento conforme a duração do vídeo também foi discutida na literatura. McGowan et al. [12] observaram uma correlação negativa entre a retenção de audiência e a duração do vídeo. E Guo et al. [8] observaram que os alunos se engajam menos em exercícios que sucedem vídeos longos.

O conteúdo apresentado também tem relação direta com o engajamento do aluno. Der Sluis et al. [20] observaram que a permanência dos estudantes aumenta em trechos com muitas informações ou informações complexas. Alguns professores notaram que há pouco engajamento dos alunos em vídeos mais teóricos, sem demonstração ou resolução de exercícios [6]. Em relação a isso, Guo et al. [8] verificaram que os alunos assistem mais tutoriais múltiplas vezes, e McGowan et al. [12] notaram que os vídeos de teoria retêm menos alunos do que os vídeos de demonstração de código. Na área de informática, McGowan et al. [12] observaram ainda que videoaulas que apresentam o código pronto, com explicação de cada trecho de código pelo professor, possuem mais engajamento dos alunos quando comparados com videoaulas que exibem o professor digitando e explicando o código.

2.2 Visualização de informação

Em MOOCs, os alunos passam mais tempo assistindo vídeos [11, 16]. Isso é um dos fatores que fazem com que a interação dos estudantes com os *players* de vídeo seja uma das fontes de dados utilizadas nas pesquisas de clusterização e predição de performance dos estudantes [2, 5]. Além disso, quando os professores foram perguntados sobre o que eles gostariam de saber com base na interação dos alunos com as videoaulas [6], a maioria respondeu querer visualizar informações de acesso aos vídeos, como estatísticas de visualização de certos trechos dos vídeos, quais trechos foram pulados e quais trechos os alunos assistiram mais de uma vez. Segundo eles, essas informações podem sinalizar ao professor o que os alunos acharam relevante e sobre quais trechos os alunos tiveram mais dúvidas.

Guo e Reinecke [9] trouxeram evidências de que a interação na *timeline* dos *players* de vídeos pode ser utilizada como um indicador de performance dos estudantes. Em outras palavras, quanto maior o número de interações na *timeline*, maior a probabilidade do estudante ter uma boa performance no curso. No contexto de cursos de inglês, um padrão de comportamento observado com relação à interação com a *timeline* foi que os estudantes que não possuem fluência no idioma realizam mais ações de *play*, *pause* e *seek* do que os estudantes fluentes [18, 19].

Um resultado empírico observado por professores é que pessoas mais velhas interagem nos AVAs de forma diferente das pessoas mais novas [6]. Isso é corroborado com os resultados obtidos por Guo e Reinecke [9], que relatam que os estudantes mais velhos tendem a obter melhor performance e a realizar mais *backjumps*⁵ na *timeline* dos *players* de vídeos do que os estudantes mais novos. Além disso, eles também observaram que os países que possuem maior razão entre o número de estudantes por professor, em salas de aula presenciais, são os que correspondem aos menores índices de *backjumps*.

Podemos encontrar na literatura ferramentas de visualização de *logs* de interação (e.g., de eventos do tipo *play*, *pause* e *seek*) por trechos de uma videoaula [5]. Como evolução dessa ferramenta,

⁵Mover o cursor da *timeline* para um instante de tempo anterior ao atual.

Chen et al. [3] apresentam uma visualização para ajudar os professores a identificarem os alunos que desistiram do curso, com base em dados de acesso aos vídeos. Contudo, os professores reportaram problemas de interpretação da visualização e de usabilidade da ferramenta.

Encontramos um único trabalho para identificar padrões de interação no *player* de vídeo que resultam em uma boa performance no curso [2]. O trabalho apresenta uma ferramenta que agrupa (em *clusters*) os alunos pelas suas interações e gera uma visualização que relaciona os *clusters* com as notas obtidas pelos alunos. Entretanto, foi observado que *clusters* que apresentam o mesmo padrão de interação possuem alunos com performances satisfatória e insatisfatória no curso. Além disso, quando a ferramenta foi testada pelos professores, eles reportaram problemas no entendimento da visualização, mesmo depois de terem passado por um tutorial.

3 PROPOSTA DE TRABALHO

Como ressaltado na seção anterior, há uma lacuna de ferramentas para auxiliar os professores na análise de *logs* de interação dos alunos. Portanto, este trabalho apresenta uma proposta de ferramenta para análise de interações com os *players* de vídeo, chamada *VideoViz*. A utilização de técnicas de agrupamento (*clustering*) de dados na *VideoViz* pode fornecer *feedbacks* ao autor do conteúdo sobre o comportamento dos usuários, assim como gerar *insights* para o autor sobre quais os trechos do vídeo são relevantes e qual o melhor dia da semana para postar conteúdo, levando em consideração critérios como: quantidade de acessos e perfil de usuários. No caso de autores, que geralmente são professores, a ferramenta pode apresentar *insights* sobre trechos do vídeo em que ocorrem dúvidas e comparar a interação nos *players* de vídeo entre os alunos com performance satisfatória e os com baixo rendimento no curso.

A *VideoViz* pode ser utilizada tanto em AVAs como em repositório de vídeos (RVs). RVs como os fornecidos pela RNP (e.g., Video@RNP⁶ e Videoaula@RNP⁷), e AVAs como os mantidos pela UNA-SUS/UFMA⁸, permitem que usuários acessem conteúdos de diversas áreas (e.g., humanas, exatas, saúde). Ao todo, ambos os RVs da RNP possuem cerca de 27 mil vídeos, com mais de 20 mil acessos por mês. Já o AVA da UNA-SUS/UFMA registra mais de 66 mil acessos por mês nos seus mais de 800 recursos utilizados nos cursos (e.g., vídeos, *ebooks*, *slides*).

Entretanto, RVs como VideoRNP e Videoaula@RNP têm uma limitação para visualização de informação com base nos *logs* de interação, pois os usuários podem acessar os vídeos sem terem feito *login*. Logo, as visualizações para esses RVs não podem correlacionar interações com o perfil dos usuários (e.g., idade, escolaridade) ou performance (i.e., notas) deles nos cursos que utilizam seus vídeos. O que é possível nos RVs é associar as interações de cada usuário com o identificador de sessão do *browser* e apresentar visualizações sobre o acesso ao vídeo (e.g., ações, data e região). Já em AVAs como o da UNA-SUS/UFMA, além de visualizar as mesmas informações de RVs, é possível gerar *clusters* de perfil, escolaridade e performance do aluno. Podemos observar essas informações resumidas na Tabela 1.

Visualização de <i>clusters</i> de interações por:	RVs	AVAs
sessão do <i>browser</i>	x	x
data (e.g., dias da semana, mês, semestre)	x	x
região de origem (e.g., cidades, estado, país)	x	x
perfil (e.g., idade, escolaridade)		x
performance em curso (e.g., notas)		x

Tabela 1: Possibilidades de análise de *logs* em RVs e AVAs

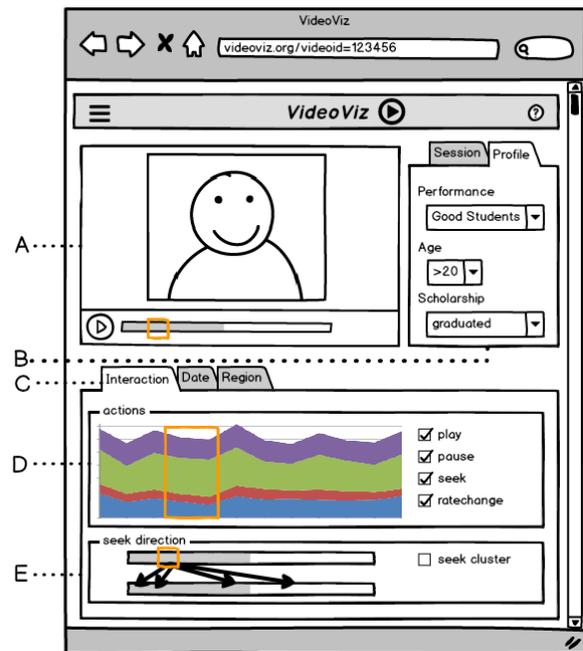


Figura 1: Esboço da ferramenta *VideoViz*

A *VideoViz* visa a apresentar para RVs e AVAs informações de padrões de acesso e de interação nos vídeos, por meio de técnicas de visualização de dados. Entretanto, as visualizações são diferenciadas dada a origem do vídeo (RVs ou AVAs). A Figura 1 ilustra um *mockup* da ferramenta, o qual detalhamos a seguir: (A) área do *player* que reproduz o vídeo, cuja navegação é utilizada como filtro de *logs* a serem analisados, ou seja, o instante de tempo do vídeo é utilizado como referência nas demais visualizações; (B) painel de indicação de como o usuário da ferramenta deseja que a visualização seja correlacionada com dados de perfil como performance, idade e escolaridade. Caso seja uma RV, essas correlações são limitadas à sessão do *browser*; (C) seletor de diferentes visualizações, presentes tanto em RVs quanto em AVAs. Nesse caso, a ferramenta apresenta opções de visualização sobre as interações no *player*. Outros tipos de visualização são por data e região de onde o usuário está acessando; (D) mostra um gráfico de área da distribuição das ações (e.g., *play*, *pause*, *seek*, *ratechange*) executadas pelo usuário no *player* durante a execução do vídeo; (E) exibe a visualização dos instantes de tempo de destino da maioria das ações de *seek* realizadas pelos usuários. Em particular, nessa visualização, a ferramenta também apresenta *clusters* de movimentação do marcador de tempo do vídeo. Isso é útil para indicar trechos repetidamente assistidos, que podem indicar

⁶<http://video.rnp.br>

⁷<http://videoaula.rnp.br>

⁸<http://avaunasus.ufma.br>

trechos com mais relevância ou de dúvidas. Além disso, a *cluster* também pode indicar trechos que são frequentemente pulados, que podem indicar trechos do vídeo a serem resumidos ou removidos.

4 CONCLUSÃO

Este trabalho apresenta uma visão geral do estado da arte de análise dos *logs* de interação em *players* de vídeos. Com o objetivo de preencher a lacuna de ferramentas para auxílio na análise desses *logs*, apresentamos o *VideoViz* como sendo uma proposta de ferramenta para ser utilizada por AVAs e RVs. Vale ressaltar que, apesar da ferramenta considerar apenas *logs* de interação em vídeos, é possível estendê-la para analisar *logs* de outros recursos (e.g., fórum, *quiz*, *ebooks*) e criar modelos de predição de abandono e performance do aluno.

A proposta faz parte de uma pesquisa em andamento. Por isso, existe a necessidade de refinamentos. Como trabalho futuro propomos a utilização de técnicas de design participativo para inclusão de autores de conteúdo (e.g., professor) no centro do processo de desenvolvimento da ferramenta. Além disso, buscamos investigar se há melhora no desempenho da turma quando os professores utilizam ferramentas de *clustering* e predição de desempenho dos alunos.

AGRADECIMENTOS

Nós agradecemos a instituição UNA-SUS/UFMA pela colaboração e parceira, e a Dalai Ribero pelas discussões sobre esse trabalho. Além disso, o presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Agradecemos também a CNPq pelo financiamento (projeto #309828/2015-5).

REFERÊNCIAS

- [1] Miguel L. Bote-Lorenzo and Eduardo Gómez-Sánchez. 2017. Predicting the decrease of engagement indicators in a MOOC. In *Proceedings of the Seventh International Conference on Learning Analytics & Knowledge (LAK '17)*. ACM, New York, NY, USA, 143–147. <https://doi.org/10.1145/3027385.3027387>
- [2] Qing Chen, Yuanzhe Chen, Dongyu Liu, Conglei Shi, Yingcai Wu, and Huamin Qu. 2016. PeakVizor: Visual Analytics of Peaks in Video Clickstreams from Massive Open Online Courses. *IEEE Transactions on Visualization and Computer Graphics* 22, 10 (oct 2016), 2315–2330. <https://doi.org/10.1109/TVCG.2015.2505305>
- [3] Yuanzhe Chen, Qing Chen, Mingqian Zhao, Sebastien Boyer, Kalyan Veeramachaneni, and Huamin Qu. 2017. DropoutSeer: Visualizing learning patterns in Massive Open Online Courses for dropout reasoning and prediction. *2016 IEEE Conference on Visual Analytics Science and Technology, VAST 2016 - Proceedings* (2017), 111–120. <https://doi.org/10.1109/VAST.2016.7883517>
- [4] Konstantinos Chorianopoulos. 2018. A Taxonomy of Asynchronous Instructional Video Styles. *The International Review of Research in Open and Distributed Learning* 19, 1 (feb 2018). <https://doi.org/10.19173/irrodl.v19i1.2920>
- [5] Conglei Shi, Siwei Fu, Qing Chen, and Huamin Qu. 2015. VisMOOC: Visualizing video clickstream data from Massive Open Online Courses. In *2015 IEEE Pacific Visualization Symposium (PacificVis)*, Vol. 2015-July. IEEE, 159–166. <https://doi.org/10.1109/PACIFICVIS.2015.7156373>
- [6] André Luiz de Brandão Damasceno, Dalai dos Santos Ribeiro, Sérgio Colcher, and Simone D. J. Barbosa. 2018. Entrevistas e mapeamento sistemático sobre análise de *logs* de interação de alunos em ambientes virtuais de aprendizagem. (2018), 11 pages. ftp://ftp.inf.puc-rio.br/pub/docs/techreports/18_08_damasceno.pdf
- [7] Ashish Dutt, Maizatul Akmar Ismail, and Tutut Herawan. 2017. A Systematic Review on Educational Data Mining. *IEEE Access* 5 (2017), 15991–16005. <https://doi.org/10.1109/ACCESS.2017.2654247>
- [8] Philip J. Guo, Juho Kim, and Rob Rubin. 2014. How video production affects student engagement: an empirical study of MOOC videos. In *Proceedings of the first ACM conference on Learning @ scale conference (L@S '14)*. ACM Press, 41–50. <https://doi.org/10.1145/2556325.2566239> arXiv:arXiv:1011.1669v3
- [9] Philip J. Guo and Katharina Reinecke. 2014. Demographic differences in how students navigate through MOOCs. In *Proceedings of the first ACM conference on Learning @ scale conference (L@S '14)*. ACM Press, New York, New York, USA, 21–30. <https://doi.org/10.1145/2556325.2566247> arXiv:arXiv:1011.1669v3
- [10] René F Kizilcec, Chris Piech, and Emily Schneider. 2013. Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge (LAK '13)*. ACM Press, New York, New York, USA, 170. <https://doi.org/10.1145/2460296.2460330>
- [11] Vitomir Kovanović, Srećko Joksimović, Dragan Gašević, James Owers, Anne-Marie Scott, and Amy Woodgate. 2016. Profiling MOOC Course Returners: How Does Student Behavior Change Between Two Course Enrollments?. In *Proceedings of the Third (2016) ACM Conference on Learning @ Scale (L@S '16)*. ACM Press, New York, New York, USA, 269–272. <https://doi.org/10.1145/2876034.2893431>
- [12] Aidan McGowan, Philip Hanna, and Neil Anderson. 2016. Teaching Programming: Understanding Lecture Capture YouTube Analytics. In *Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education (ITICSE '16)*. ACM Press, New York, New York, USA, 35–40. <https://doi.org/10.1145/2899415.2899421>
- [13] Jiezhong Qiu, Jie Tang, Tracy Xiao Liu, Jie Gong, Chenhui Zhang, Qian Zhang, and Yufei Xue. 2016. Modeling and Predicting Learning Behavior in MOOCs. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining (WSDM '16)*. ACM, New York, NY, USA, 93–102. <https://doi.org/10.1145/2835776.2835842>
- [14] Cristóbal Romero and Sebastián Ventura. 2010. Educational Data Mining: A Review of the State of the Art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 40, 6 (Nov 2010), 601–618. <https://doi.org/10.1109/TSMCC.2010.2053532>
- [15] Terezinha Saraiva. 1996. Educação a distância no Brasil: lições da história. *Em Aberto* 16, 70 (abr 1996), 17–27. <https://doi.org/10.24109/2176-6673.emaberto.16i70.2076>
- [16] Daniel T. Seaton, Yoav Bergner, Isaac Chuang, Piotr Mitros, and David E Pritchard. 2014. Who does what in a massive open online course? *Commun. ACM* 57, 4 (apr 2014), 58–65. <https://doi.org/10.1145/2500876>
- [17] George Siemens and Ryan S. J. d. Baker. 2012. Learning Analytics and Educational Data Mining: Towards Communication and Collaboration. In *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge (LAK '12)*. ACM, New York, NY, USA, 252–254. <https://doi.org/10.1145/2330601.2330661>
- [18] Selen Türkay, Hadas Eidelman, Yigal Rosen, Daniel Seaton, Glenn Lopez, and Jacob Whitehill. 2017. Getting to Know English Language Learners in MOOCs: Their Motivations, Behaviors, and Outcomes. In *Proceedings of the Fourth (2017) ACM Conference on Learning @ Scale (L@S '17)*. ACM, New York, NY, USA, 209–212. <https://doi.org/10.1145/3051457.3053987>
- [19] Judith Uchidiuno, Jessica Hammer, Evelyn Yarzebinski, Kenneth R. Koedinger, and Amy Ogan. 2017. Characterizing ELL Students' Behavior During MOOC Videos Using Content Type. In *Proceedings of the Fourth (2017) ACM Conference on Learning @ Scale (L@S '17)*. ACM, New York, NY, USA, 185–188. <https://doi.org/10.1145/3051457.3053981>
- [20] Frans Van der Sluis, Jasper Ginn, and Tim Van der Zee. 2016. Explaining Student Behavior at Scale: The Influence of Video Complexity on Student Dwelling Time. In *Proceedings of the Third (2016) ACM Conference on Learning @ Scale (L@S '16)*. ACM Press, New York, New York, USA, 51–60. <https://doi.org/10.1145/2876034.2876051>
- [21] Hongyan Wang, Xingwei Hao, Wenjiang Jiao, and Xingwu Jia. 2016. Causal Association Analysis Algorithm for MOOC Learning Behavior and Learning Effect. In *2016 IEEE 14th Intl Conf on Dependable, Autonomous and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*. IEEE, 202–206. <https://doi.org/10.1109/DASC-PiCom-DataCom-CyberSciTec.2016.53>

Mapping and review of audiovisual editors and streaming platforms based on collaborative technologies in real time

Brett Stalbaum
Speculative Design
University of California, San
Diego (UCSD)
9500 Gilman Drive,
La Jolla, CA
stalbaum@ucsd.edu

Cicero I. da Silva
Brazil Telehealth Networks
Federal University of São Paulo
(Unifesp)
Rua Sena Madureira, 1500
4o andar, São Paulo, SP
cicero.silva@unifesp.br

Jane de Almeida
PPG Education, Art and
History of Culture
Mackenzie University
Rua da Consolação, 930,
São Paulo, SP
janeddealmeida@post.harvard.edu

Maria Amelia Eliseo
Department of Computing and Informatics
Mackenzie University
Rua da Consolação, 930,
São Paulo, SP
mamelia@gmail.com

Vic Von Poser
Mf Art Digital
University of the Arts (UAL)
45-65 Peckham Road, London,
SE5 8UF
totalvic@gmail.com

ABSTRACT

The main goal of this article is to map and review real time video streaming collaborative softwares, in order to analyze platforms that are also able to edit films. Real time high resolution film transmission (4k, 8k or above) results in the excess of data generated and, consequently, in a high value for investment in editing, storage and content distribution features. The retrieval and storage of the data (content) becomes complex and expensive. This article aims to design a software review for the use of systems that are applicable to the needs directly related to the digital education field, considering that nowadays the video is an important tool to support teaching and learning processes and needs a better structuring from the educational managers of communication and information technologies.

KEYWORDS

Video processing, video editors, streaming

1 INTRODUCTION

This section shows the current scenario regarding editor tools, highlighting their properties related to the ideal features for an on-line video editor with streaming. Among the existing tools with possibilities of streaming on the web, the best known and most

popular are:

- a) YouTube; b) Livestream; c) Justin TV; d) XSplit Broadcaster;
- e) Kaltura; f) Clipbudget; g) Bambuser; h) Wirecast; i) Open Cast

Some stream tools are only dedicated to video features, others stream and store the content for later viewing and editing, if the user wants to edit the material in post-production. Softwares who have video editors allow users to insert clips, mark the beginning and end of a video, and crop unwanted materials. Users can even add a soundtrack.

Below is a brief description of the most popular streaming tools, some of them with editing properties:

- a) YouTube is the most widely used streaming model today. The platform allows the editing of the material transmitted after the streaming is finished. However, it does not allow consecutive editing of the transmitted content, i.e. it is not possible to insert audio, cut or even change any content consecutively to the streaming. To carry out a streaming on YouTube you need specific tools that make the connection between the camera, the microphone and the compression system that Youtube will use. For the sake of simplicity, the XSplit Broadcaster tool stands out for its screen capture, streaming and addition of plugins that help in post-production, as well as being more intuitive and already adapted for Youtube use. Another option is to use Quik (G + Hangout on Air), which performs content streaming through the use of the Hangout tool. When the Hangout transmission is finalized, the content can be made available on Youtube.
- b) In addition to this tool, there is the Kaltura software, which has both its commercial and its open source

version with several plug-in options for post-production and web application modules, such as integration with the CMS (Custom Management System) WordPress (VideoPress), with the Joomla CMS, among others.

c) Wirecast Go. The tool enables real-time transmission through mobile devices such as tablets and cell phones.

The characteristics of Wirecast are:

- c.1) a single camera (front or rear of mobile devices) live;
- c.2) is capable of streaming directly to YouTube with just one click;
- c.3) add images, logos and graphics to your photo library stream;
- c.4) basic editing (resizing and moving graphics);
- c.5) unlimited streaming;

d) Skype Qik: video sharing application that allows you to send audio content to your closest friends on your Windows Phone, iPhone or Android;

e) Bambuser: this platform is a good option to use with mobile technologies (cell phones and tablets).

Among the video capture software, a fair option is the Open Broadcaster Software, which is free and interoperable with other systems, such as Kaltura, Adobe Connect and Red5. It is one of the most used systems by gamers to produce their live game streaming via Youtube and Vimeo, among others.

The streaming software system is important because it allows the captured content to be sent through other platforms. In addition, these softwares perform the encoding of the video and audio signal, creating a specific video file that can be transmitted through these platforms, such as Youtube, Vimeo, Wevideo, among others. In systems that generate video display protocols, each tool usually generates its own viewer, which makes it easy to display the video stream both on Internet browsers and on the screens of mobile applications. This view is common in applications such as Bambuser, Qik etc. where you can view real-time streaming of content streamed through your mobile (mobile or tablet) applications. In other video systems, such as Youtube, it is possible to send different video streams from the most varied sources, since the compaction and reading systems of these viewers do not contemplate the reading of many formats and, in this sense, it is necessary a software that "reads" the content and presents it. However, these softwares do not allow users to view in real time a timeline with frames being accumulated one after another. The streaming tools usually do not generate separate audio and video files, only allowing post editing of the content. Another lack of this kind of software is that there is no buffering, which would allow the user to decide what they would like to save and to "publish" in real time from the streaming.

2 SOFTWARES FOR STREAMING

2.1 Real time systems

The table below shows a comparative table between two online video platforms and the respective feeds of video streams

supported by these platforms, which means that it is not possible to separate a video stream feed from the Bambuser application, for example, and broadcast it in real time at the same time on Youtube, i.e. Youtube, in this case, accepts a variety of video feeds, not getting dependent only on their own structures of transmission of video feeds.

Software	Formats	Players
Bambuser	MP4	Bambuser
Youtube	MP4, FLV, MPEG, MOV, entre outros	Youtube, XSplit Broadcaster, Quick (G+ Hangout)

Below is a brief list of the most popular streaming tools, some of them with editing properties:

- a. Wowza¹
- b. Wirecast²
- c. Vjdirector³
- d. WSE Manager⁴
- e. Open Cast⁵
- f. 23⁶

3 VIDEO EDITORS: STATE OF THE ART

As soon as there is no real time editing software platform, currently the process of assembling or editing an audiovisual product consists of selecting, ordering, and adjusting the plans or moments in order to achieve the expected result. In this sense, this section reviews the top three most used, by it's market share, video editor softwares and points out the lack of a feature that can be, in the near future, crucial for professionals that work based on live streaming of contents. Before the digital age, one can compare the editing process with writing a text on a typewriter. Before typing, the text should be well outlined, drafted and revised; any subsequent changes to the text on paper were complicated, time-consuming and costly. So were editions of audiovisual products, in which the final product as a whole was thought before going into the post-production process. At the beginning of the cinema, these were physical processes like cutting and gluing of films. With the evolution of technology, recordings began to use magnetic tapes (BETAMAX, VHS), to select and re-record, which resulted in loss of material quality. With these processes there was no opening for error or chance, everything had to be thought out and planned in advance. To this

¹ See <https://www.wowza.com/>

² See <http://www.telestream.net/wirecast/overview.htm>

³ See <http://nagashare.com/vjdirector/>

⁴ See <http://wsemanager.com/>

⁵ See <http://www.opencast.org/>

⁶ See <http://www.23video.com/api/>

type of edition we call "linear editing". With the advent of the computer, "editing software" works with hardwares (Video Cards, HDS), including versions for personal computers. The first system adopted by the market was the hardware and software mix created by NewTek Video Toaster for the Amiga 2000 computer, as shown in Figure 1. The mechanism presented an interface inspired by analog processes, such as the scissor icon representing the cutting tool of the old movie theater moviolas. At the time, this represented a breakthrough in the mobility of images within the video, such as easier erasing or replacing images. It also consisted of a series of effects and tools, transitions and enhancement of low cost images that were once complicated to accomplish. Large concepts presented by this mechanism still resists in modern editing software, such as digital cutting tools and the Timeline concept, as seen in the interface shown in Figure 2. However, it was no longer necessary to think of an integral final product and many experiments can be performed by inserting and removing images within the editing line. This would be the beginning of the concept of "non-linear editing".



Figure 1: Software and hardware system for video editing in the 1990s, proposed by NewTek Video Toaster.



Figure 2: System interface proposed by NewTek Video Toaster 2.

3.1 Apple Final Cut

In the professional universe it is considered the software that has the more sophisticated features and requires more time and practice to be manipulated. At the beginning of the digital age, one of the software preferred by independent audio-visual professionals was Apple Final Cut 7 (Mac OS X only, released in

2009). It presented better features and a user-friendly interface, with the possibility of greater interaction and manipulation of the material. In 2011, Apple released an update of this software called Final Cut X. In this new version, the software, which was the most used by professionals in the field, has an interface and a range of effects more similar to amateur software mentioned above. Figure 3 shows the interfaces of Final Cut Pro and Final Cut Pro X.



Figure 3: On the left, the interface of Final Cut Pro; on the right, the interface of Final Cut Pro X.

3.2 Adobe Premiere Pro

Many users, however, did not adapt to the new presentation of the Final Cut software, because in an attempt to make the software friendly and easy, Apple ended up launching a very similar program with the editing software dedicated to beginners, and in that sense, the professionals finished considering that this version had many limitations. Users were gradually migrating and looking for other software that suited their needs and had a more familiar interface than they were accustomed to. The most sought after was Adobe Premiere Pro (Figure 4), which, with its 2014 update, started to present better and more user-friendly features for professional users, and even improved communication with other Adobe software such as Adobe After Effects, Photoshop and Illustrator. It was also a great advance in the dynamicity of the possibility of importing diverse materials, of various formats and codecs, that operated together in a same timeline, making better use of the time for the publisher, who no longer had to import the videos into the timeline every time they wanted to edit them, or even make a selection of a particular scene.



Figure 4: Adobe Premiere Pro interface.

3.3 Avid

In the medium and large video production companies, we can highlight the use of the Avid Media Composer tool (Figure 5). Avid is a digital software that still works with the concept of a linear edition. It was released in 1991 as a set of hardware and software, and evolved into a widely used and respected software, especially in large works and by editors/directors accustomed to the beginning of video editing. One advantage of this software is that it generates its own media, so it does not work with the original raw material, which can be copied and saved for other purposes. But, this software is still complicated in terms of a non-linear video editing, presenting problems with mobility and takes changes. To work with this software, you need to have the thought of the movie as a whole in advance again, as in the analogic film editing process.

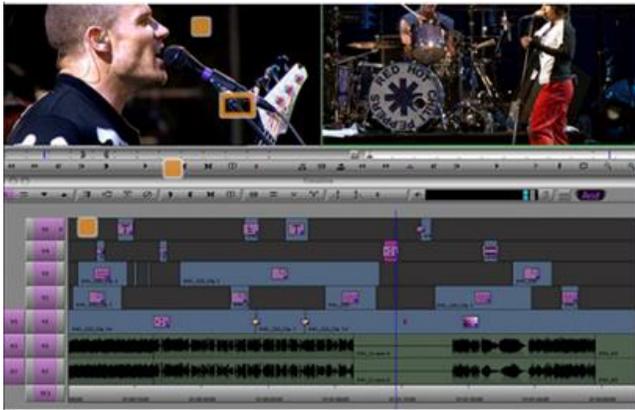


Figure 5: Avid Media Composer interface.

4 RESULTS AND DISCUSSION

If the user is a professional or amateur, linear or non-linear, pre-created or editable timelines and digital x analogic editing issues matters more than we could expect. By choosing editing softwares, among the many available, users will always depend on the goals, interests and intents to cut, paste, and manipulate information to build a new meaning for the recorded and scattered images in the digital world. It is important to stress that right now no edition platform can edit online videos in real time, creating a shortage of options for streaming video editors that work in the education and in entertainment fields. On August 30th, 2018 PC Mag, one of the most respected magazines in the technology field, ran an article describing the most important video editors based in their market share.

Below is a comparative table of video editors and its features published by PC Mag that corroborates our article review:

Product	Adobe Premiere Pro CC	Corel VideoStudio Ultimate	CyberLink PowerDirector	Pinnacle Studio Ultimate	Adobe Premiere Elements	Magix Movie Edit Pro Premium	Wondershare Filmora	Nero Video 2017	Apple Final Cut Pro X	Apple iMovie (for Mac OS X)
Lowest Price	SEE IT	SEE IT	SEE IT	SEE IT	SEE IT	SEE IT	SEE IT	SEE IT	SEE IT	SEE IT
Editors' Rating	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Best For	Professionals	Enthusiasts	Enthusiasts	Prosumers	Enthusiasts	Enthusiasts	Consumers	Enthusiasts	Professionals, Prosumers	Consumers
Number of Video Tracks	Unlimited	20	100	Unlimited	Unlimited	99	11	Unlimited	Unlimited	2
Motion Tracking	—	✓	✓	✓	✓	✓	—	—	—	—
Multicam Editing	✓	✓	✓	✓	—	—	—	—	✓	—
3D Editing	✓	✓	✓	✓	—	—	—	—	—	—
Supports 360° VR Content	✓	✓	✓	✓	—	✓	—	—	✓	—
Keyword Tag Media	—	—	—	✓	✓	—	—	—	✓	✓
Supports 4K XAVC-S Format	✓	✓	✓	✓	✓	✓	—	—	✓	✓
Exports to H.265 (HEVC)	✓	✓	✓	✓	✓	✓	—	—	✓	—

Figure 6: PC Mag table with video editors softwares. [7]

ACKNOWLEDGMENTS

This work was partially supported by The São Paulo Research Foundation (FAPESP) under its BPE Program (FAPESP grant 17/20434-6).

REFERENCES

- [1] Heide Alvarez. AmLight Consortium Research & Education Network Helps Transmit FIFA World Cup in 8K from Brazil to Japan. Center for Internet Augmented Research and Assessment (CIARA) - Florida International University. Miami-USA. 22 jul 2014. Available online: <http://ciara.fiu.edu/images/release_8K-WorkdCup_network.pdf>. Access: Feb 06, 2017.
- [2] Jane de Almeida; Maria Amelia Eliseio; Cicero I. da Silva; Helena Prates; Vic Von Poser; Brett Stalbaum and Nilton Furtado. Sensemaking: A Proposal for a Real-Time on the Fly Video Streaming Platform. Creative Education, v. 07, 2016,p.2515-2523.
- [3] Jane de Almeida; Brett Stalbaum; Alfredo Suppia and Cicero I. da Silva. Passages on Brazilian scientific cinema. Public Understanding of Science, 2017, Vol. 26(5) 579–595. DOI: <https://doi.org/10.1177/0963662516683638>.
- [4] Jane de Almeida; Guido Lemos Souza Filho; Carlos Batista; Lucenildo Aquino Junior; Jose Ivan Vilarouca Filho; Manoel Neto; Cicero I. da Silva; Leandro Ciuffo; Iara Machado and Clayton Silva. Integration of Fogo Player and SAGE (Scalable Adaptive Graphics Environment) for 8K UHD Video Exhibition. Journal of Computer and Communications, v. 02, 2014, p. 50-55.7.
- [5] Jane de Almeida; Cicero I. da Silva and Thiago Andre (Orgs.). CineGrid: futuros cinematócos. 1. Ed. São Paulo: CINUSP, 2017.
- [6] Gilberto Sobrinho. Cinemas em redes. Tecnologia, estética e política na era digital. Campinas, Papirus, v.1, 2016.
- [7] Michael Muchmore. The best video editing software of 2018. PC Mag. Accessed September 1st, 2018. Available online: <https://www.pcmag.com/article2/0,2817,2397215,00.asp>

Sistema de videocolaboração para o museu interativo do TCE-PB baseado no modelo DA

Amanda Azevedo
Design Audiovisual/LAVID/UFPB
João Pessoa, Brasil
amanda.azevedo@lavid.ufpb.br

Sandro Gonçalves
Design Audiovisual/LAVID/UFPB
João Pessoa, Brasil
sandro.goncalves@lavid.ufpb.br

Valdecir Becker
Design Audiovisual/LAVID/CI/UFPB
João Pessoa, Brasil
valdecir@ci.ufpb.br

Raoni Kulesza
LAVID/CI/UFPB
João Pessoa, Brasil
raoni@lavid.ufpb.br

Fábia Carolino
TCE-PB
João Pessoa, Brasil
fluna@tce.pb.gov.br

Agnaldo Macedo Filho
TCE-PB
João Pessoa, Brasil
amfilho@tce.pb.gov.br

ABSTRACT

This work describes a system focused on interactive audiovisual experiences developed for the State of Paraíba Court of Auditors (TCE-PB). The system includes professional and amateur productions to enable social control using multimedia technologies. As a result, the video collaboration solution, based on the Audiovisual Design model, is presented.

KEYWORDS

Audiovisual Design, Social Control, Fruition.

1 INTRODUÇÃO

O Museu Interativo do Tribunal de Contas do Estado da Paraíba TCE-PB é um projeto com o objetivo de conscientizar a população sobre o controle social, que usa tecnologias da informação e comunicação. Para tanto, a produção de conteúdos audiovisuais, dispostos em ambientes multimídias, compõe uma estratégia de disseminação de uma cultura de fiscalização. Esse projeto é uma parceria do Grupo de Pesquisa Design Audiovisual com o Laboratório de Aplicações em Vídeo Digital (LAVID) e o Tribunal de Contas do Estado da Paraíba (TCE-PB). O Projeto é composto por aplicativos, games, ambientes de realidade aumentada e conteúdos audiovisuais. Este artigo descreve os sistemas de fruição de produções profissionais e colaborativas do componente audiovisual do Museu, chamado Cine-TCE [1]. O escopo deste módulo inclui as produções audiovisuais interativas do projeto sendo apresentados os sistemas de Gerenciador de Conteúdos no

portal web e três ambientes de fruição e visualização: o Mapa Colaborativo no modo web, que exhibe todos os conteúdos enviados, a Mesa Interativa em painéis sensíveis ao toque, encontrada no Museu, e um sistema de Apoio a Palestras, onde os conteúdos disponíveis podem ser incluídos em apresentações no Auditório do TCE-PB. São contempladas duas formas de entrada de conteúdos: vídeos institucionais, produzidos pelo LAVID, que retrata a atuação do TCE-PB; e vídeos de curta duração produzidos por estudantes da rede pública do estado da Paraíba, que mostra as dificuldades de execução em projetos públicos e sugere soluções para tais problemas.

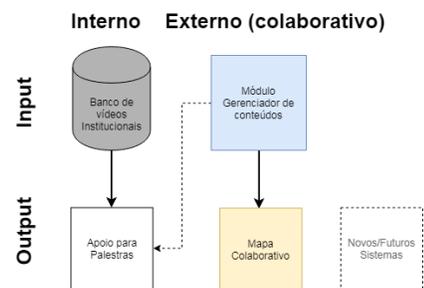


Figura 1. Relação entre conteúdos e formas de visualização.

Conforme descrito na figura 1, cada entrada de conteúdo tem uma abordagem de visualização direta. Considera-se também nesta proposta a possibilidade de utilização de dados do Módulo de Gerenciamento de Conteúdos para desenvolver funções ou atividades dentro do Sistema de Apoio a Palestras bem como futuros sistemas.

O Museu busca tornar a visitação contínua potencializando o programa já existente TCE Escola e Cidadania, no qual os estudantes tem acesso ao acervo do Memorial e a palestras no auditório do TCE-PB, sobre a função na auditoria das contas públicas, combate a corrupção e controle social. O projeto foca na rede estadual de ensino, tanto na produção dos conteúdos, quanto

nas visitas ao Museu. Para tanto, pretende-se usar os recursos do Museu para conscientizar os estudantes e professores sobre a importância do controle social. Conteúdos como cidadania, políticas públicas e convivência em sociedade, que já fazem parte do dia a dia dos estudantes, servirão de base para a disseminação da cultura de fiscalização.

2 MATERIAIS E MÉTODOS

Para desenvolver os cenários de fruição, o Design Audiovisual foi utilizado como modelo teórico e método. O Design Audiovisual é um modelo conceitual e um processo metodológico para a criação e avaliação de conteúdos audiovisuais, mediados por interfaces digitais, integrando de forma interdisciplinar os conceitos e processos dos estudos de mídia e Interação Humano Computador. Aplicações e análises deste modelo foram realizadas em estudos anteriores de fruição audiovisual [5] [6] [7] [8] [9]. O modelo descreve quatro papéis que os indivíduos assumem ao consumir conteúdos audiovisuais: Audiência, Sintetizador, Modificador e Produtor. O papel da Audiência representa o comportamento “passivo” do indivíduo diante da fruição, como a visualização dos conteúdos. As ações do Sintetizador são: compilar, classificar, comentar, recomendar e compartilhar o conteúdo, construindo, dessa forma, uma “identidade” digital, um perfil encenado em uma rede social. O Modificador domina softwares para manipular e recriar conteúdos, ampliando a noção de engajamento para a de apropriação. O papel do Produtor são os responsáveis pelo design e criação dos conteúdos, seja de forma autônoma ou em grupos de mídia. Cada papel tem seus níveis elevados, descritos como Players, quando o indivíduo se envolve e usa os recursos disponíveis da interface para fruir um conteúdo audiovisual. As mudanças de papéis ocorrem quando há um aumento do engajamento e da interação do indivíduo com os sistemas, quando há a percepção da qualidade do conteúdo junto a utilidade do software.

Os sistemas apresentados promove uma experiência direcionada à facilitação da visualização, filtragem e personalização de conteúdos para envolver o indivíduo nos níveis de engajamento apresentado nos papéis de Audiência. As etapas apresentadas explanam desde a idealização do projeto até a implementação das soluções. A criação da proposta, as possibilidades de comunicação e interação, e a realização da pesquisa de usuário com alunos das escolas públicas seguiram as Linhas de Design do modelo de Design Audiovisual. Finalmente, o projeto foi formalizado e implementado usando métodos ágeis da engenharia de software.

3 CENÁRIOS DE TESTES

Nos tópicos seguintes descreveremos como cada sistema supre as demandas do Museu Interativo do TCE-PB, sendo elaborado por etapas, seguido dos cenários de testes em telas de navegação.

Etapa 1: Delimitamos um público específico para imergir na proposta do projeto. Alunos, professores e palestrantes foram os indivíduos guias para o desenvolvimento dos sistemas e estratégias de produção de conteúdos audiovisuais.

Etapa 2: A produção de um portal web do projeto que inclui um Gerenciador de Conteúdos, servindo como ferramenta para entrada dos dados através do envio dos vídeos, e concentra os conteúdos que são acessados pelos demais sistemas. No portal o professor realiza o cadastro e interage com as funcionalidades de gerenciamento referentes aos envios, seleções e alterações dos vídeos dos alunos.

Etapa 3: O sistema do Mapa Colaborativo corresponde à visualização dos conteúdos, sendo alimentado pelos vídeos enviados através do Gerenciador. O sistema tem uma página à parte do portal, contendo o mapa da Paraíba, em que cada município é um ponto de acesso aos respectivos conteúdos dos alunos.

Etapa 4: O segundo sistema de visualização é uma representação pelo mesmo Mapa Colaborativo web, mas acessível por uma Mesa Interativa touchscreen e um projetor de curta distância, dispostos dentro do Museu. Esse sistema permite, além da visualização, a criação de playlists para exibição.

Etapa 5: O terceiro sistema é o de personalização de palestras, dedicados exclusivamente aos funcionários do TCE-PB que ministram essas apresentações durante a visita das escolas no auditório. O sistema oferece ao palestrante formas de mapear e escolher os vídeos compondo uma apresentação mais atrativa para os visitantes.

3.1 Inputs de Interação

Diante destas aplicações descreveremos detalhadamente como ocorrerá o funcionamento técnico, a interação com as interfaces e a relação com os conteúdos audiovisuais do Museu com base nos *inputs* de fruição dos sistemas.

3.1.1 Gerenciador de Conteúdos

Os casos de uso neste sistema são a realização do cadastro pelo professor, que preenche um formulário com dados pessoais, fornecendo também sua matrícula da rede de ensino para que o cadastro seja autenticado no sistema.

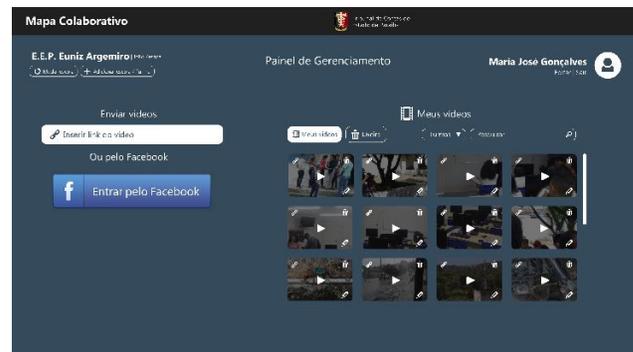


Figura 2: Parte da interface do sistema de Gerenciamento de Conteúdos

Realizado o cadastro e o login, o professor é guiado à interface principal do Gerenciador, onde pode adicionar os vídeos via desktop, autorizar os conteúdos dos alunos usando as postagens dos perfis do Facebook captados por API Graph, ativar as permissões para a publicação do conteúdo na plataforma ou, ainda, inserir o link do vídeo para que os conteúdos sejam direcionados ao Mapa Colaborativo. Os requisitos funcionais do sistema são: login, upload, remoção e edição de vídeos, filtro por escolas e turmas participantes.

3.1.2 Mapa Colaborativo

Os conteúdos enviados pelo Gerenciador de Conteúdos servirá como *input* para o Mapa Colaborativo. A visualização no mapa

pode ser acessada no modo web e na Mesa Interativa disposta no Museu.



Figura 3: Parte da interface do sistema do Mapa Colaborativo.

Os requisitos para o Mapa Colaborativo no modo web são: o acesso pelo local de envio, regiões de auditoria do TCE-PB e na galeria de vídeos pela busca das escolas e acesso ao player de vídeo.

3.2 Outputs de Fruição

As saídas dos sistemas do TCE-PB corresponde a personalização de vídeos, mesclando autorias dos vídeos profissionais e dos colaborativos, que promove experiências com os conteúdos de forma inédita nas visitas ao Museu.

3.2.1 Mesa Interativa

Para a interação na Mesa são acrescentados os requisitos de criação de playlists de exibição. Os casos de uso para a Mesa Interativa são as sessões de exibição no Museu que permite a visualização de forma coletiva. Como requisitos desejáveis será necessário inserir uma página de login para o professor, que guiará as sessões junto aos alunos no Museu, possa elaborar uma personalização de playlists com os vídeos, que poderão gerar ao final da sessão sua propagação através do compartilhamento nas redes sociais usando uma URL ou QR Code.

3.2.2 Apoio a Palestras

O sistema de Apoio a Palestras é de uso dos palestrantes para criar playlists de vídeos institucionais e colaborativos, de acordo com o objetivo da apresentação. Esse sistema permite a escolha de trechos ou dos vídeos completos. Os requisitos deste sistema são a criação de apresentações, mapeamento dos conteúdos, ajustes, disponibilização de apresentações já realizadas que ficam registradas no sistema e player de vídeos acessível por aplicativo mobile, para controle da apresentação. Os requisitos desejáveis deste sistema será o acesso a URL ou QR Code da apresentação personalizada do palestrante para que os alunos possam compartilhar em suas redes sociais, que ao identificar suas autorias torna-se um fator atrativo para estimular a propagação.



Figura 4. Parte da interface do sistema de Apoio a Palestras.

Todo o sistema está instalado localmente, com conexão entre dispositivos feita por rede local. A realização de testes focou nos requisitos de login e envio de vídeos como etapas iniciais de implementação para a primeira versão da aplicação.

4 CENÁRIOS DE PRODUÇÃO E CONSUMO PELAS LINHAS DE DESIGN AUDIOVISUAL

O uso integrado do conteúdo e interface em cenários multimídias resulta em experiências de fruição cada vez mais complexas. Conforme apresentado no estudo sobre arquitetura de design colaborativo [6], o indivíduo ao explorar o conteúdo se relaciona com a interface em momentos passivos de visualização e momentos ativos de interação e propagação. Os papéis de Audiência do Design Audiovisual são frutos de ações baseadas nas quatro linhas conceituais de Design: Identidade, Motivação, Experiência e Conteúdo que visam uma experiência focada nos fatores humanos do indivíduo, gerando novas formas de transmissão e recepção de conteúdos, diferente como encontrado no modelo Broadcast [5].

Demonstramos, portanto como ocorre o percurso de interação que conduz as experiências de fruição entre indivíduo e sistemas, por meio das demandas de produção e consumo.

4.1 Linha do Conteúdo

A linha do conteúdo é definida pelas práticas de produção audiovisual, em que o estudante que produz conteúdo assume o papel do Produtor, de criador da obra, desde a concepção estética à criação do conteúdo.

4.2 Linha da Experiência

Esta linha foca no envolvimento, nas curiosidades e desafios que o indivíduo realiza para fruir o conteúdo, elevando suas ações a níveis extremos, que corresponde a características de mais de um papel e emerge o Player. Cada percurso de interação promove experiências diferenciadas. O uso do Mapa Colaborativo proporciona o indivíduo a visualizar o conteúdo, configurando-se Audiência, e possibilita alterar um momento passivo para ativo ao se relacionar com suas funcionalidades como: selecionar vídeos usando as ferramentas de busca tornando-se Audiência-Player, apoiar e comentar os vídeos gerando competições, atuando como Sintetizador-Player.

4.3 Linha da Identidade

As ações nesta linha são geradas por preferências e personalizações que o indivíduo realiza no processo de fruição. A

experiência no Museu através da Mesa Interativa gera alternância entre as ações de Audiência (visualização) para Modificador (novas possibilidades de conteúdo), que ao elaborar playlists organizadas em categorias de temas, escolas ou vídeos mais apoiados, alteram o sentido inicial e seus contextos, gerando um novo produto. No caso de uso do sistema de Apoio a Palestras os papéis de Audiência e Sintetizador estão mais evidentes entre os estudantes, que ao assistir a palestra atua como Audiência e ao receber o vídeo personalizado compartilhando nas redes sociais torna-se um Sintetizador, incorporando suas opiniões, críticas e configurando um “perfil” de conteúdo a ser exposto nas redes sociais.

4.4 Linha da Motivação

No processo de fruição o indivíduo possui uma intenção, que busca atender as suas necessidades intrínsecas e extrínsecas. Identificamos os principais casos de uso focados na motivação. O professor recebe os conteúdos disponibilizados pelos alunos, submetendo ao Gerenciador de Conteúdos, tornando-se assim Sintetizador. A mediação é focada na motivação para a produção destinada aos os alunos, bem como selecionar os conteúdos que devem representar a escola, sendo indispensável sua participação.

O palestrante no momento que realiza a seleção dos vídeos para gerar uma nova apresentação atua como Modificador, no qual busca despertar o engajamento da Audiência para o que pretende comunicar na palestra.

O foco das linhas de Design é direcionar a experiência para a mudança dos papéis, que não necessariamente são lineares, sendo baseado em conjuntos, tornando-se dinâmicos. A experiência de fruição pode ocorrer em dois ou mais papéis simultaneamente: o indivíduo que assiste uma obra (Audiência), participa de discussões nas redes sociais (Sintetizador) e interfere no conteúdo ao remixar a produção inserindo suas impressões e gerando novos sentidos (Modificador ou Produtor).

5 RESULTADOS

Os sistemas do museu tem as linhas de Design como modelagem do produto, fazendo uso de *Affordances* que foram pensadas para as ações de produção, personalização e propagação de conteúdo audiovisual em um ambiente midiático, de acordo com a proposta do museu. As *Media Affordances* [7] tem como objetivo guiar o indivíduo através da correta percepção associadas ao seu contexto no ambiente de fruição que envolve percepções físicas (*affordances* que podem ser ativamente manipuladas, com usos potenciais facilmente perceptíveis, compostas por tecnologias usadas para fruição de conteúdos), gráficas (*affordances* presentes em interfaces gráficas interacionais implementadas por software e responsáveis por parte da mediação entre indivíduos e conteúdo) e simbólicas (*affordances* presentes na narrativa ou inseridas no conteúdo audiovisual, onde as *affordances* físicas são mais sutis e as relações cognitivas e sensoriais dependem não apenas de tecnologia e interfaces, mas também da compreensão de elementos narrativos). Com base nas *Media Affordances* são estabelecidos Gatilhos de Interação (Interaction Triggers - IT), divididos em Ação (Triggers of Action - ToA) e Inércia (Triggers of Inertia - ToI) que possibilitam ao indivíduo desempenhar a função dos papéis de Audiência. Com base nesses parâmetros serão avaliados as produções dos alunos e a os níveis de engajamento identificando se atinge ou não os objetivos

necessários para cada papel, que desempenham a videocolaboração dentro do sistema.

6 CONSIDERAÇÕES FINAIS

Essas apropriações dos conteúdos e integração no uso dos sistemas apontam para novas perspectivas de autorias, personalização e modificação em sistemas colaborativos, gerando produtos multimídias inovadores. Como perspectiva futura deste projeto aponta-se as novas opções de fruição descritas como requisitos desejáveis. Esse tipo de propagação permitirá que se atinja um maior número de pessoas através das estratégias apresentadas. Diante deste compartilhamento surgirão novas implicações do conhecimento acerca do Museu e, conseqüentemente, no uso dos demais sistemas.

7 AGRADECIMENTOS

Os autores gostariam de agradecer a Marianna Teixeira, Jaqueline Donin e Thiago Vasconcelos, cujas pesquisas foram fundamentais para a viabilização da presente proposta.

8 REFERÊNCIAS

- [1] CHAUDHRY, A. N. TOSCANO, R. M. N. AZEVEDO, A. S. BECKER, V. Pesquisa documental e propostas de identidade visual e de interação. Módulo WP7 Cine do projeto Espaço Interativo do TCE, abril, 2018.
- [2] CHAUDHRY, A. N. TOSCANO, R. M. N. AZEVEDO, A. S. SOUZA, H. DONIN, J. BECKER, V. Pesquisa documental, propostas e protótipos para vídeo e interação. Módulo WP7 Cine do projeto Espaço Interativo do TCE, maio, 2018.
- [3] CHAUDHRY, A. N. TOSCANO, R. M. N. AZEVEDO, A. S. SOUZA, H. BECKER, V. Pesquisa e produção de conteúdo. Módulo WP7 Cine do projeto Espaço Interativo do TCE, junho, 2018.
- [4] TOSCANO, R. M. N. AZEVEDO, A. S. GONÇALVES, S. DONIN, J. TEIXEIRA, M. VASCONCELOS, T. CHAUDHRY, A. N. BECKER, V. Produção de vídeos institucionais, visualização e interação com o conteúdo. Módulo WP7 Cine do projeto Espaço Interativo do TCE, julho, 2018.
- [5] BECKER, V. GAMBARO, D. RAMOS, T. Audiovisual Design and the Convergence Between HCI and Audience Studies. Springer International Publishing AG 2017.
- [6] TOSCANO, R. M. N. BECKER, V. SILVEIRA, L. COUTINHO, S. BURGOS, L. Arquitetura de design colaborativo para imersão temporal e espacial em vídeos de altíssimas resoluções e HFR. O futuro da videocolaboração: perspectivas. In S. B. de Computação (Ed.) Gramado, RS, 2017.
- [7] BECKER, V. GAMBARO, D. RAMOS, TOSCANO, R.M.N Audiovisual Design: Introducing “Media Affordances” as a Relevant Concept for the Development of a New Communication Model. Applications and Usability of Interactive Television. 6th Iberoamerican Conference, JAUTI. Aveiro, Portugal, October, 2017.
- [8] BECKER, V., GAMBARO, D., RAMOS, T. S., & BEZERRA, E. P. (2018). Design Audiovisual: a interseção dos estudos de audiência com a Interação Humano Computador. *Conexão-Comunicação e Cultura*, 17(33).
- [9] GAMBARO, D., BECKER, V., RAMOS, T. S., & TOSCANO, R. (2018). The Development of Individuals’ Competencies as a Meaningful Process of the Audiovisual Design Methodology. In International Conference on Human-Computer Interaction (pp. 68-81). Springer, Cham.

Videoconferência multiambientes: o sistema Multipresença com foco na área da saúde

Valter Roesler
Instituto de Informática, UFRGS –
Universidade Federal do RGS
Porto Alegre, RS, Brasil
roesler@inf.ufrgs.br

Guilherme Longoni
Mconf Tecnologia Ltda
Porto Alegre, RS, Brasil
guilherme@mconf.com

Rafael Valle
RNP – Rede Nacional de Ensino e
Pesquisa
Rio de Janeiro, RJ, Brasil
rafael.valle@rnp.br

ABSTRACT

This paper presents the Multipresence system with a focus on its uses for eHealth. The Multipresence system performs multi-technology videoconferencing, that is, it allows the interoperation of various technologies and devices such as: Telepresence rooms in high definition (until Full-HD); Ultra-telepresence room (4K); Sharing of content (images, texts, applications); Legacy videoconferencing systems (SIP standard); software application installed on the personal computer; Web conferencing (via web browser); Mobile devices and SIP phones. UFRGS University works in the coordination of the technological development of the Working Group GT-Multipresença in partnership with RNP (National Research and Education Network) and with Mconf Tecnologia Ltda. The project is in its fourth year of development, with functional prototypes very stable, and already been used for transmission and interaction in robotic surgeries. The objective of this article is to disseminate this technology as a driving force in the creation of video collaboration panels and integration within and between countries, in the educational or e-Health area.

KEYWORDS

Multimedia; e-Health; Distance Learning; web conference; video collaboration.

1 INTRODUÇÃO

A comunicação através de videoconferência é algo comum nos dias de hoje, permitindo a interação remota de pessoas, economizando tempo e dinheiro. Atualmente existem diversas formas de videoconferência, como a Telepresença, os Sistemas de Sala e a Webconferência, entre outros, como detalhado em [4].

Algumas questões que ainda são tópicos de pesquisa e desenvolvimento atualmente envolvem a interoperabilidade entre os diversos tipos de videoconferência, a escalabilidade da solução e a adaptabilidade à banda em redes heterogêneas. Alinhado com essa necessidade, está sendo desenvolvido um

sistema de videoconferência denominado “Multipresença”. O objetivo inicial do sistema é a interoperação transparente de pelo menos o seguinte:

1. Sala de telepresença em alta definição (Full HD – 1920x1080p).
2. Sala de ultra-telepresença em ultra alta definição (UHD 4K – 3840x2160).
3. Acesso através de sistemas de videoconferência de sala (padrão SIP, integrando com Polycom, Cisco, e outros).
4. Acesso em alta definição através de programa aplicativo no computador pessoal.
5. Acesso através de webconferência.
6. Acesso através de dispositivos móveis.
7. Compartilhamento de conteúdo.

Uma facilidade do sistema de Multipresença, além do modelo de integração explicado acima e do baixo custo, é a possibilidade de adaptar o espaço físico de acordo com as necessidades do usuário, através da troca de layouts (ou visões do ambiente).

A interoperação das múltiplas tecnologias propostas pelo sistema Multipresença permite diferentes aplicações compartilharem a mesma sala, como por exemplo: a) aulas presenciais e a distância; b) dinâmica de grupo local e remota; c) reuniões em U; d) reuniões de diretoria; entre outras. Cada aplicação tem suas próprias demandas de comunicação em termos de qualidade, interoperação e colaboração.

O sistema Multipresença foi desenvolvido no escopo do Programa de Grupos de Trabalho da Diretoria de Pesquisa e Desenvolvimento da Rede Nacional de Ensino e Pesquisa (RNP). Desde 2002, o programa fomenta o desenvolvimento de novos produtos e serviços para a comunidade de ensino e pesquisa brasileira.

A seção 2 apresenta de forma resumida o sistema. A seção 3 apresenta alguns usos especificamente em e-Saúde, e a seção 4 apresenta a rede e as considerações finais.

2 O SISTEMA MULTIPRESENÇA

Em agosto de 2015 um protótipo da solução foi apresentado durante o evento “Forum RNP”, em Brasília, Brasil. A Figura 1 mostra uma imagem do painel. As quatro TVs da esquerda mostram o compartilhamento de conteúdo através do SAGE2 (*Scalable Amplified Group Environment*, explicado a seguir), onde havia compartilhamento de logos, imagens e aplicações entre

In: V Workshop “O Futuro da Videocolaboração” (WCT-Video 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pôsteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.
©2018 SBC – Sociedade Brasileira de Computação.
ISBN: 978-85-7669-435-9.

Brasília e Porto Alegre (2.000km de distância). As duas TVs do canto inferior direito mostram a videoconferência em alta definição ponto a ponto entre Brasília e Porto Alegre. As duas TVs do canto superior direito mostram a integração com o sistema de webconferência Mconf, com a TV da esquerda mostrando dois vídeos (um usuário conectado via tablet e outro via notebook) e a TV da direita mostrando slides de uma apresentação. A TV maior no canto direito da imagem mostra uma transmissão em tempo real em ultra-alta definição (4K) de Porto Alegre, onde um arquivo YUV de 3Gbit/s (simulando uma câmera ao vivo) era comprimido e enviado para Brasília em tempo real a aproximadamente 25 Mbit/s.



Figura 1: Foto do painel durante o Fórum RNP 2015.

Ambos locais (Brasília e Porto Alegre) estavam conectados através do *backbone* da RNP, suportando toda a banda requerida para todas essas comunicações em tempo real.

Há um vídeo de demonstração do sistema Multipresença no endereço <https://www.youtube.com/watch?v=lud2KwzstCM>. Este vídeo foi gravado durante o evento em agosto de 2015.

A comunicação com alta definição, incluindo a integração com sistemas legados de videoconferência, é efetuada através de um módulo chamado “PRAV Player”.

O Mconf [4, 5, 6, 7] é o módulo de webconferência da solução do Multipresença. Ele é responsável pela integração com dispositivos móveis, bem como pela integração com telefones SIP.

O módulo de compartilhamento de conteúdo foi criado através de uma integração do Multipresença com o sistema de código aberto de colaboração chamado SAGE2 [1], desenvolvido pelas universidades de Illinois e do Havá.

O SAGE2 é uma ferramenta que permite aos usuários compartilhar suas telas, arquivos e aplicativos através de redes IP, criando um campo virtual onde participantes remotos e locais podem compartilhar e visualizar os conteúdos [2].

A Figura 2 mostra o painel instalado na UFRGS durante uma reunião. O painel possui 6 TVs. As 3 TVs de baixo estão configuradas para alta definição, e estão se comunicando respectivamente com o INCA (Instituto Nacional do Câncer) no Rio de Janeiro, com a Escola Bahiana de Medicina e Saúde na Bahia, e com a Escola Superior de Redes da RNP em Brasília. As duas TVs do canto superior esquerdo mostram os participantes conectados através de webconferência (no caso são 7 participantes mais o vídeo da sala aonde a foto foi tirada,

portanto 8 vídeos), e a TV do canto superior direito mostra a apresentação de slides utilizada na reunião.



Figura 2: Painel do Multipresença instalado na UFRGS.

O impacto do sistema para a sociedade é grande, pois o Multipresença é um sistema de baixo custo e possui aplicações para diversas áreas como e-Saúde e teleeducação, permitindo a criação de salas multiuso de fácil reconfiguração, bem como transmissão de cirurgias e tele consultoria médica. A seguir alguns exemplos de uso do sistema em e-Saúde mostrando o potencial dessa TIC.

3 CASOS DE USO EM E-SAÚDE

Caso 1: Em 2016, foi utilizado uma versão do sistema Multipresença focado em transmissão de cirurgias para efetuar a transmissão do VII SIVA (Simpósio Internacional de Videocirurgia Avançada), um evento organizado a cada dois anos na Irmandade da Santa Casa de Misericórdia de Porto Alegre (ISCOMPA – <http://www.santacasa.tche.br>), um grande complexo de 7 hospitais, localizado em Porto Alegre, RS. O software é intitulado MIR (Multimedia Integrated Room) [3].

O objetivo do simpósio, que tem a duração de 3 dias, é a discussão de técnicas cirúrgicas, com a transmissão simultânea de 3 salas cirúrgicas ao vivo para um teatro remoto, e debates sobre os procedimentos. Cirurgiões de várias partes do Brasil e do mundo participam do evento. Duas cirurgias são exibidas no teatro através de dois projetores Full-HD. Normalmente a terceira sala está em preparação.

A Figura 3 apresenta uma visão do teatro do CHC (Centro Histórico Cultural) durante o evento. Pode-se perceber as duas transmissões acontecendo simultaneamente, sendo que o operador pode escolher, via tablet de controle, quais transmissões devem chegar ao palco. As possibilidades são várias, como a interna e externa de qualquer sala, a interna de uma sala e a externa de outra, e assim por diante. É possível também apresentar a imagem interna e externa no mesmo projetor, dividindo as telas em até 8 imagens simultâneas. Esse recurso foi utilizado em alguns momentos para mostrar várias imagens de duas salas, mas quando possível, passava-se para tela cheia a fim de ter um maior detalhamento da imagem.

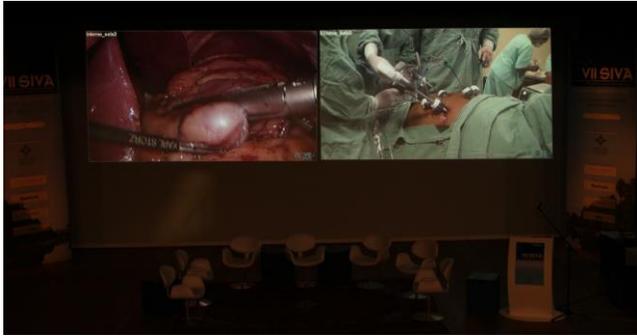


Figura 3: Visão do teatro do CHC com a transmissão acontecendo em tempo real.

A interação entre as dezenas de cirurgiões participantes no teatro e o cirurgião na sala cirúrgica, executando e debatendo sobre sua técnica, foi bastante proveitosa. Em vários momentos houve debates calorosos sobre as técnicas utilizadas, mostrando que a tecnologia estava inserida de forma transparente e apoiando a comunicação dessa aplicação de e-Saúde. Isso só foi possível porque o atraso do sistema é bastante baixo, na ordem de 125ms, similar à sensação do ser humano numa conversa presencial. Além disso, a qualidade da transmissão permitia a visualização do procedimento de forma muito melhor do que estando na própria sala.

Nesse caso não foi necessária a integração com webconferência ou sistemas legados. A transmissão foi efetuada somente entre hospitais do mesmo complexo, utilizando a rede de computadores interna (via IP).

Caso 2: O Instituto Nacional do Câncer (<http://www2.inca.gov.br>) está atualmente utilizando o sistema Multipresença para transmissões de cirurgias robóticas em alta-definição e via web.

As cirurgias robóticas são efetuadas com o robô Da Vinci, que permite ao cirurgião controlar os braços mecânicos (pinças e outros instrumentos cirúrgicos) a partir de um console que apresenta as imagens internas do paciente.

As cirurgias com o robô permitem fazer uma incisão menor, aumentando também a precisão, praticamente zerando o tremor do médico e alcançando cavidades que com a videolaparoscopia tradicional não é possível. Para os pacientes, o robô permite cirurgias menos invasivas e, por isso, menos traumáticas, com redução do tempo de internação e recuperação mais rápida.

Apesar de ainda ser um projeto piloto, a ideia do instituto é criar uma rotina de transmissões, com objetivo de disseminar conhecimento em técnicas cirúrgicas.

A Figura 4 mostra uma transmissão de cirurgia em tempo real realizada em 2017. As duas telas de baixo mostram a cirurgia em tempo real (visão interna e externa do paciente), enquanto a TV de cima mostra algumas pessoas conectadas através de webconferência.



Figura 4. Transmissão de cirurgia integrando alta definição e webconferência.

A Figura 5 mostra a mesma transmissão sendo visualizada através do sistema de webconferência. As pessoas que assistem via navegador web possuem as mesmas imagens de quem está com o painel, porém a qualidade da transmissão é inferior, devido ao fato do acesso ser adaptada para baixa banda.

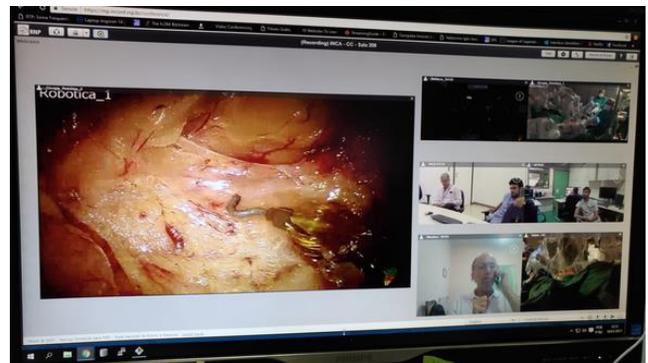


Figura 5: Visão da cirurgia através de webconferência.

Caso 3: O sistema Multipresença está sendo utilizado na parte de videoconferência do projeto Teleoftalmo – Olhar Gaúcho, inaugurado dia 10 de julho de 2017. O projeto permite telediagnósticos em oftalmologia, e prevê oito consultórios remotos no Estado do Rio Grande do Sul, com um centro de controle em Porto Alegre onde os médicos analisam diversos aspectos em relação à visão do paciente.

O projeto foi liderado pelo TelessaúdeRS-UFRGS com o patrocínio do Hospital Moinhos de Vento no âmbito do programa PROADI-SUS do Ministério da Saúde e da Secretaria estadual de Saúde do Rio Grande do Sul. Também estão envolvidos o Ministério da Saúde e prefeituras dos municípios.

O oftalmologista necessita da imagem em alta definição para verificar diversas questões relativas ao olho, como foco, acuidade visual, entre outras. O paciente sai do pólo remoto já com diagnóstico completo e até receita de óculos estará disponível com o médico solicitante, caso paciente necessite.



Figura 6: Ambiente durante atendimento.

A Figura 7 mostra uma imagem no Centro de Controle localizado no Telessaúde-RS. Cada médico oftalmologista deverá controlar intercaladamente duas salas, otimizando seu tempo e agilizando o atendimento. A figura mostra parte do ambiente de controle, onde a médica verifica os dados dos instrumentos remotos e também interage através da videoconferência.

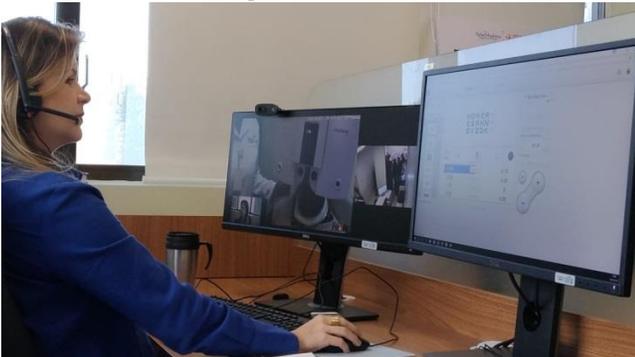


Figura 7: Ambiente durante atendimento.

Essa inovação vem reduzindo a espera por atendimento na especialidade de oftalmologia, uma das mais procuradas pela comunidade. A fila de pacientes aguardando uma consulta era, originalmente, na ordem de 9 mil pessoas que levariam mais de um ano para serem atendidas e diagnosticadas. O projeto está diminuindo drasticamente esse problema, com atendimentos remotos de mais de mil pacientes por mês.

4 CONSIDERAÇÕES FINAIS

O sistema de Multipresença está implantado em aproximadamente 15 instituições no Brasil. Todas as instituições já possuem pelo menos um painel do projeto.

A inovação promovida pelo sistema Multipresença é a modularidade e ubiquidade oferecida, que não se vê em sistemas similares. Assim, comunicações 4K estariam interagindo com salas de telepresença em Full HD, com sistemas de sala, com computadores stand-alone, com computadores via videoconferência em redes de menor velocidade e também com dispositivos móveis. A sala pode ser reconfigurada rapidamente e

facilmente para acomodar diferentes ambientes de colaboração, otimizando espaço físico e permitindo a melhor utilização dos recursos.

O custo do sistema é baixo, visto que a instituição necessita basicamente de um kit com algumas TVs (ou monitores ou um projetor), e um computador genérico ligado a essas TVs. Os outros servidores podem ser mantidos pela operadora de serviços (NREN nacional).

Considera-se de extrema importância que o desenvolvimento de uma nova TIC seja feito de forma interdisciplinar, com participantes ativos do lado da medicina (médicos, pacientes, cirurgiões), do lado da informática (programadores e arquitetos de software) e eventualmente das áreas humanas (psicólogos, por exemplo). Com especialistas em várias áreas idealizando um projeto comum, este fica muito mais útil para a sociedade.

A rede formada pelo projeto Multipresença envolve, do lado tecnológico, a presença da UFRGS (líder do projeto), da empresa Mconf Tecnologia (que apoia tecnicamente e patrocina parte da equipe) e da RNP (que oferece uma visão gerencial de uso do sistema e patrocina parte do desenvolvimento).

AGRADECIMENTOS

Este trabalho foi patrocinado pela RNP (Rede Nacional de Ensino e Pesquisa), a NREN brasileira, e pela Mconf Tecnologia. O software de alta definição foi patrocinado pela FINEP (Financiadora de Estudos e Projetos).

REFERÊNCIAS

- [1] MARRINAN, T. et. al. SAGE2: A New Approach for Data Intensive Collaboration Using Scalable Resolution Shared Displays. In Proceedings of the 10th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing. 2014.
- [2] SAGE (2015). Electronic Visualization Laboratory, University of Illinois at Chicago; Laboratory for Advanced Visualization & Applications, University of Hawai'i at Manoa. SAGEnext: Next Generation Integrated Persistent Visualization and Collaboration Services for Global Cyberinfrastructure. Si2 PI Workshop Forum, National Science Foundation. February 17, 2015.
- [3] ROESLER, V. et. al. MIR: A Low Cost Digital Operating Room. In IEEE Healthcom 2014, Natal, RN, Brazil. 16th International Conference on E-health, Networking, Application & Services. 2014.
- [4] ROESLER, V.; CECAGNO, F.; DARONCO, L. C.; DIXON, F. Mconf: an open source multiconference system for web and mobile devices. Book Chapter In: Multimedia - A Multidisciplinary Approach to Complex Issues, ISBN 978-953-51-0216-8, 1ed., 2012, p. 203-228.
- [5] ROESLER, V.; CECAGNO, F.; DARONCO, L. C.; MARINS, A. Mconf: sistema de multiconferência escalável e interoperável web e dispositivos móveis. In: Tecnologias da Informação e Comunicação na América Latina, 2012, Lima. TICAL 2012.
- [6] ROESLER, V.; CECAGNO, F.; Daronco, Leonardo C.; DIXON, F. Mconf: towards a global webconference system. In: TERENA Networking Conference, 2012, Reykjavik. TNC 2012.
- [7] ROESLER, V.; CECAGNO, F.; DARONCO, L. C.; MARINS, A.; STANTON, M. Interoperable multi conferencing technology as a basis for an open, global web conferencing network. In: UbuntuNet-Connect, 2013, Kigali, Rwanda. 2013.

Índice de Autores

Aguiar, Yuska Paola C.	73	Guedes, Alan L.V.	65, 171, 175
Albuquerque, Monck Charles N. de ...	69	Husemann, Ronaldo	127
Almeida, André	97	Kulesza, Raoni	183
Almeida, Fernanda Kécia de	69	Leal Filho, Durval	167
Almeida, Jane de	179	Lima, Guilherme F.	65
Almeida, João Paulo D. de	19	Lima, Manuella A. C.B.	73
Almeida, Rodrigo L.A.	91	Longoni, Guilherme	187
Alves, Romero C.	163	Lopes, Frederico	97
Amaral, Leandro A. do	139	Macedo Filho, Agnaldo	183
Anderle, Daniel Fernando	81	Macedo, Márcio C.F.	49
Andrade, Rossana M.C.	91	Maia, Marcio	115
Araripe, Alexandre	115	Martins, Francisco R.F.	69
Araujo, Carlos V.S.	37	Melo, Guilherme Afonso M.S.	151
Araújo, Italo L. de	91	Milidiú, Ruy Luiz	171
Araújo, Manuela A. de	121	Monier, Elza B.	175
Araújo, Tiago M. U. de	73	Nogueira, Rodrigo Ramos	81
Azevedo, Amanda	183	Nunes, João Victor G. de S.	65
Bandeira, Renan	115	Oliveira, Yuri S. de	69
Barbosa, Maria L.K.	133	Peixoto, João Paulo J.	103
Barbosa, Simone D.J.	175	Pimentel, Maria da Graça C.	145
Barrére, Eduardo	31	Poser, Vic von	179
Becker, Valdecir	163, 167, 183	Rabêlo, Dilson J.L.	175
Bittar, Thiago J.	139	Reis, Luana S.	73
Busson, Antonio J.G.	171	Rodrigues, Kamila R.H.	145
Calumby, Rodrigo T.	53	Roesler, Valter	127, 133, 187
Carolino, Fábria	183	Sales, Angelina S.S.	73
Carvalho, Camilo A.	121	Santiago, José Marcelo S.	69
Carvalho, Luis Paulo da S.	121	Santos, Gabriel N.P. dos	171
Cazella, Sílvio César	133	Santos, Marília G.	49
Chaves, Samir	77	Scalco, Leonardo F.	145
Colcher, Sérgio	65, 171, 175	Sette, Ioram S.	25
Costa Filho, Roberto I.T. da	127	Silva, Alberto V.D. da	13
Costa, Daniel G.	103	Silva, Cicero I. da	179
Costa, Marcia	133	Silva, Glayson B.S.	163
Costa, Wdson	121	Silva, Paulo Henrique T.	121
Cristo, Marco A.P.	37	Silva, Romulo	57
Cunha, Bruna C.R.	145	Silva, Wanderson B. da	53
Damasceno, André Luiz de B.	175	Soares Neto, Carlos de S.	151, 171
Dantas, Paulo Filipe	57	Soares, Eduardo R.	31
Darin, Ticianne G.R.	91	Soares, Letícia P.	121
Durão, Frederico A.	7, 19	Sousa Silva, Diogo V. de	7
Eliseo, Maria Amelia	179	Souza, Antonio C.S.	49
Ferraz, Carlos André G.	25	Souza, Helder Bruno A.M.	167
Fortes, Renata P. de M.	139	Souza, Vítor E.S.	109
Gasparoni Júnior, Mário	127	Stalbaum, Brett	179
Giusti, Rafael	37	Tabosa, Davi B.	61
Gomes, João	115	Teixeira, Marianna C.	167
Gonçalves, Sandro	167, 183	Toscano, Rafael M.	167

Tozzi, Tatiana	81	Xavier, Marcus Rafael	25
Trinta, Fernando	77, 115	Zaine, Isabela	145
Valle, Rafael	127, 187	Zancan, Luciano	133
Viana, Windson	57, 61	Zemor, Mario F.	133
Vieira, Vaninha	13	Zupeli, Breno L.	109
Viel, Caio C.	145		
Vilela, Allyson	97		