I. INTRODUCTION

Brazil will receive a lot of tourists during the megaevents that it will host in the next years. In order to optimize the host services to this public, Brazil’s Ministry of Tourism created a profile of these tourists. The study was conducted during Africa’s World Cup, in order to identify potential market pathways [13]. According to this research, the tourists were mostly European and North-American, 83% being men, 60% single men and 86% graduates from colleges. Each tourist stayed for an average of 15 days during the event and spent R$11,400 (no airfares included) [21].

Another important result is the fact that 80% of those tourists have never come to Brazil, which allows for the conclusion that those megaevents are an opportunity to receive an international group of tourists that is different from the usual one because this public is different and would not come to Brazil otherwise.

This represents a bigger challenge to academics and entrepreneurs, because showing to those tourists a receptive country, with socially and economically sustainable activities and innovations in service may cause the megaevent to become a multiplier of the future number of tourists, further developing the country.

Considering that Brazil exports qualified Computer Science workers, we believe that academics, together with businessmen, may effectively create a computational legacy that is interesting and may be used also after the megaevents.

In order to offer information, products and services that are both personalized and assertive to persons/clients/tourists, the Brazilian and Worldwide academic community has used Recommendation Systems (RS), which may be created and applied to different domains, such as movie theatre and film recommendations.

A. Scenario

Since its creations, the movie industry has created a large variety of movies and for every new launch there are several pieces of information made available to the general public (managers, moviegoers, etc). This information on the movies are not treated or filtered before arriving to the clients, hence are not very accessible and personalized for every moviegoer (the prospective customer of this industry).

Information on movies are usually made available in a way that causes information overload and makes it more difficult for the client to make a personalized choice. This information overload is an old concern [13] and treating this problem has been one of the challenges of the Recommendation Systems [20]. The RS indicate items of potential interest for users. The insertion of contextual information such as movie time, exhibition place, ticket price and age recommendation help the recommendation process, refining it. With this information at hand, it is possible to predict items according to the context in which the user is inserted [31].

Besides using context information, it is possible to use personality traits from each user in order to improve the recommendation process. According to Nunes [24] [25][48], using personality traits in Recommendation Systems improves the recommendation offered.

Creating an application that aggregates technologies and services available in mobile devices may help gather context information such as: the movie theatre closer to the user, the movies that are being shown at this particular movie theatre, etc. Using Global Positioning Systems (GPS) together with the

Keywords—Recommendation Systems, Personality, contextual information, Megaevents.
resources of recommendation systems may help find context information, allow for refinement of the recommendation and for quicker and more effective availability of the information.

This paper is organized as follows. Section 2 shows some characteristics of recommendation systems and their techniques. Section 3 discusses related work. Section 4 presents the case study in the scenario described in this section followed by the model and prototype details. Experimentation and discussed are presented in section 5, while section 5 presents the conclusions and future works. Finally, section 7 we conclude with all the references used in this paper.

II. RECOMMENDATION SYSTEMS

Usually, moviegoers base their choices on the recommendations of friends, expert opinion or even on other sources, such as social media. Recommendation systems help and potentialize this natural social process that already exists among people [27].

For Resnick and Varian [31] the biggest challenge is to find the relationships among the user’s interests, in order to forecast the items that will interest them the most.

In order to forecast items, it is necessary to find data on them. The gathering process can be either explicit or implicit, as follows:

(i) Gadinho and Lhuillier [9] say that the explicit gathering on products is the one in which the user needs to manually infer his preferences on items. The problem with this methodology is that the user may not specify his interests fully, allowing for masking of possible results.

(ii) The implicit data gathering process uses techniques of content extraction and text mining among others in order to gather data from the client without his awareness. In this model information on possible interest areas is extracted and applying association techniques it is possible to suggest services and products [9]. The implicit gathering also helps in getting user data forming or improving his profile [24].

According to Cazella, Nunes e Reategui [4], as the information on the user grows, so does the pertinence of the recommendation on products, services and/or people. In order to translate the psychological factor of an individual with the goal of creating his profile, it is necessary to define his personality [10]. To do that, there is a very interesting approach that is the use of personality traits that allow for the psychological differentiation between persons using measurable and definable traits [33] [47].

After gathering data, we apply the techniques to recommend those collected items. In the session that follows, we present the main recommendation techniques.

Recommendation Techniques

Information filtering techniques interact dynamically with the users and at each new action performed the algorithm must seek for a new item that corresponds to the preferences of that user at that given moment and the search does not have to be initiated by the user himself.

The main recommendation techniques are the following:

1. Filtering based on content (FBC): This technique is based exclusively on the variables User and Item, comparing the content of the users’ preferences to the contents of each item. Item categorization is the big problem of this technique, because the search for information in a specific context may modify the meaning of a metadata. The process of identifying the content of the attributes and the differentiation of its meaning in that specific context is a big challenge [1][2].

2. Collaborative Filtering (FC): uses the users’ evaluations on other items and compare them with other users’ evaluations. Adomavicius and Tuzhilin [2] stress that an important characteristic of FC is that this model uses a domain independent recommendation technique. It is indicated for content recommendation that cannot be adequately described by metadata.

3. Contextual Information Based Filtering (FBIC): this came to fill the void of the tradition 2D approaches described above, which are based solely on the pair User X Item and that do not use contextual information. For this model, the context is inserted in an implicit way in all human activities, besides helping communication. When we realize the current context, we can make evaluations, decisions and adapt out behavior according to the situation [1][2][3]. For a person to make decisions in a way that is appropriate to the situation, it is necessary to understand the context [34].

4. Filtering based on Psychological Aspect (FBAP): According to González, the emotional factor is an influence on the rational thinking process when the user receives a recommendation. Hence, the author proposed filtering based on other contexts, working with psychological aspects such as emotional intelligence and social interaction [11]. Given how big the context concept is, this work proposes a change in nomenclature proposed by González. The author proposed “other contexts” and this work specializes the nomenclature in Filtering based on Psychological Aspect (FBAP) in the highest possible level and Filtering Based on Emotions (FBE) to define the concept of context when it is connected to emotional characteristics of an individual. Soon after, Nunes [24] proposed an approach inside Filtering based on Psychological Aspect (FBAP) including personality aspects in the filtering process, which we name Filtering Based on Personality (FBP).

5. Nunes [24] defined Recommendation Systems based on personality, which were afterwards defined by Hu and Pu [13] as PBRS (Personality-based Recommender System). In order to capture the essence of the individual personality differences for each individual, it is usually applied a test to understand his personality. The test is an empirical research that is able to reveal a specific set of personality traits that differentiate one individual from the other.
6. Hybrid Filtering (FH) uses one or more recommendation techniques, usually combining a collaborative approach based on content with collaborative filtering. FH has been widely used in Recommendation Systems because they use more than one recommendation technique to prevent classic cold start problems (which are defined as the problem that comes from the lack of initial information on the user that are necessary to generate an adequate and assertive recommendation for him).

The current work will use FH, because of the good characteristics inherent to that approach.

III. RELATED WORK

Several previous papers have approached different proposals for recommendation of products such as movies. NetFlix [6] proposes using data mining and content-based filtering, besides collecting information implicitly in which the system infers users’ preferences.

On the other hand, MovieLens [23] [38] [39] and IMDB [15] use FBC and explicit gathering of user information, in which he must infer his preferences on items, besides using FC techniques to improve the recommendation process.

Some works, like Jinni [17], HunchMovies [18], WhattoRent [36], in order to improve the efficiency of the RS have used implicit gathering together with explicit one, where the user manually informs his preferences using questionnaires grading items or selecting areas of personal interes.

Filtering based on contextual information (FBIC) is used in recommendation systems such as Alfred [42] [43] and DITTO [5]. There RS do not recommend movies, only services for the user using GPS data to gather contextual information.

CinemaKI [14], despite collecting contextual information from the user, can recommend movies according to an explicit collection performed by the service. Since we did not find any references that describe the technique used by CinemaKI, we performed an exploratory research on the system and inferred that it uses FC techniques.

Even though we did not find any reference that describes the technique used by Jinni, we used the system and inferred that it uses FBP by applying questionnaires for personality extraction from new users in order to know more about them [18][41]. In the questionnaire, the profile is drawn according to 12 types of filmgoers that are available at the system> the introspective, the hero, the individualist, the drama addicted, the antisocial, the extremist, the alternative reality, the idealist, the art of the fugue, the master of the mind, the strategist and the social philosopher.

According to Lawler [18], HunchMovies uses an intelligent recommendation engine called “The Living Room” that generates personalized movie recommendation using FBP and FC, including the genre that is most adequate to the user.

In order to gather information, WhattoRent [36][42] uses explicit data collection to gather information on the emotional state of the user.

At table 1 we present a comparative among the related works discussed here and the Personal_Movie model proposed at this paper. We described a few characteristics and the X in the column represents that this characteristic is present at the system represented in that row and the ? represents the fact that we cannot be sure, because of lack of references, which technique is used. The techniques evaluated were the information gathering process, data mining use, information filtering technique and whether the system uses or not contextual information, FBP, FC e FBIC.

As table 1 illustrates, the strong point of the prototype we propose is the fact that it uses FBIC with the FBC, FBP and FC techniques, because gathering contextual information will allow it to generate a recommendation that is more suitable to the user current context.

IV. CASE STUDY: PERSONAL_MOVIE

The case study proposed in this paper is a testbed for the construction of RS for megaevents.

The case study was based on movie RS, using contextual information, especially those connected to geolocalization and personality traits, together with the traditional recommendation techniques (notice that the movie is treated as a product, hence the system can be easily adapted to any other type of information, product and service to be recommended at megaevents).

New RS approaches used to recommend movies help forecast which movies will be more suitable to the current user context.

A. Personal_MovieModel

Considering the described scenario, we describe in this section the proposed model.

At Personal_Movie, contextual information allow for the discovery of the movies that are more suitable to the current user context. It performs a pre-filtering that implicitly gathers information using FBIC, when the movies currently showing are collected from the movies that are located near the user

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>RELATED WORKS X RECOMMENDATION TECHNIQUE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IMDB</td>
</tr>
<tr>
<td>Coleta Implicita</td>
<td>?</td>
</tr>
<tr>
<td>Coleta Explicita</td>
<td>x</td>
</tr>
<tr>
<td>FC</td>
<td>x</td>
</tr>
</tbody>
</table>
location (inferred through an interaction with the GPS interface). Afterwards, the collected items (movie programming data) are stored at the database (figure 1).

Personality traits are gathered based on an interaction with the system Personality Inventory [29] (figure 2) that is based on the Big Five model, in which are treated characteristics such as extroversion, sociabilization, neuroticism and openness on the factors and other characteristics in the sub-factors [46]. This information is included in Personal_Movie at the FC process in order to find similar users. Using personality traits intends to enrich the user profile.

With the information obtained, both on users and products, a similarity ration is calculated. This similarity is based on FC for products and FBAP for users, according to the characteristics manipulated by the Big Five model. The algorithms that implement the recommendation techniques (FBIC, FC, and FBAP) are stored inside the Web Service of the recommendation system, so that once generate the list of recommendable items the same are sent to the mobile device (figure 2).

The device, after receiving the list of recommendable items, execute the adjustment of the post-filtering phase (steps 6, 7 and 8 of figure 2) that is defined empirically taking into consideration the context option selected by the user. Figure 3 shows the available options for the user to select his current context.

For instance, taking into consideration a scenario where the user selected option 5, which indicates that he has something scheduled and does not have the time to watch a long movie, movies that have this duration will not be recommended. In another scenario where option 2 was selected, the movies whose age recommendation is for persons whose ages are 16-years or more will be excluded from the list of recommendable items.

B. Prototype

The prototype Personal_Movie uses the concept of SOA (Service-Oriented Architecture), which is an approach to create distributed computational systems that hide the business logic into services that can be used in a loosely coupled way [7].

Personal_Movie has a webservice that contains all the algorithms that are responsible for gathering the context information as well as generating the recommendation of geolocalized movies. The gathering process is performed through an interface of the mobile device and after that the information is sent to a HTTP server that contains the webservice to be processed in order to make the recommendation available to the user. There is also the option of the user evaluating whether or not the recommended movie is adequate to his needs and desires. This information is fed into the system, increasing the accuracy of the next suggestions (figure 4).
The recommendation engine uses the technique FBIC in order to gather information on movies that are currently showing at movie theatres and after that applies FBC to classify that movie information, storing them afterwards in a database using the SQL Server 2008 R2 DBMS [16].

In order to finalize the recommendation process, the prototype uses the FC technique to calculate the similarity among items that were evaluated by the user using the Pearson correlation metric [4].

If the user wishes to receive more specific recommendations, he only need to select the “P” option (for Personality) (figure 5) available at the interfaces “Recommendation and Programming”, and he will be taken to the personality interface, where the items will be exposed taking into consideration the user’s personality traits. The items will be forecast according to the similarity of the personality applying FBP among users, what, according to Nunes and Cazella [27] may decrease cold start problems.

During the first interactions with the system, the movies that are going to be shown at theatres and schedules closer to the current ones, since the user did not interact with any item, in order to decrease cold start problems.

Each user evaluation through Personal_Movie is sent to the server and stored in the system, what guarantees that next time the system is executed by any user, this information will be processed and may help generate recommendations different from the previously generated.

As we can see by this discussion and by analyzing figure 6, the flow that the prototype follows uses the FH technique to forecast items that are probably more assertive. Another characteristic that differentiates this system from the others depicted in section III is the availability of the forecast items through the mobile device, allowing the users to receive recommendations quickly and very effectively, helping them in their natural choice process.

V. EXPERIMENT

As stated above, this experiment is a testbed on how to use the proposed technology to personalize in an efficient way the offer of information, products and services to tourists that will come to megaevents in the next years.

The testbed measures the users perception on the movie recommendation made by the Personal_Movie system.

A. Research hypothesis

The following research hypothesis was raised to perform the experiment: Personal_Movie can extract context information and use them together with personality traits to generate a recommendation based on FBP and FBIC.

B. Sample profile

This experiment was conducted on twenty six students of Computer Science at the Federal University of Sergipe. The data set obtained represents the habits of those 26 users that watched movies during the experiment, that was conducted from April 10th to May 8th, 2012.

We made a mapping of the collected sample, as shown in figure 7, that plots the age and sex, in order to offer subsidies for posterior data analysis and a future mapping of consumed items according to age and sex.
C. Work methodology

In order to evaluate Personal_Movie we created a mobile application using GPS data to gather the context in which the user was inserted, making it viable to perform the experiment on any given population.

The research was made available at a web address where we demonstrate the steps necessary to install the application in the device as well as the system manual containing detailed information on how to use the application correctly.

The first step in the experiment was the mapping of the personality traits of each user at the Inventory [29] in an explicit gathering process. Afterwards, we imported the personality tests results into Personal_Movie, together with complementary information obtained at registration (e-mail, password, age, etc.)

After this gathering step, the users at the experiment accessed Personal_Movie and received three options to evaluate movies, as shown in Figure 8.

At the evaluation based on similarity among items were generated recommendations based on the FC, FBC and FBIC techniques. At the evaluation based on similarity among users movie recommendation was generated using the techniques FBC, FBIC and FBP.

D. Evaluation metrics

The Mean Absolute Error was used in both traditional types of evaluations based on FC and FBP, according to the following formula [13] [40]:

\[
MAE = \frac{\sum_{i=1}^{N} |F_i - O_i|}{N}
\]  

(1)

In the precision application, the feedbacks received by the sample were converted into a binary scale, being the ones from [3-5] considered relevant and the grades [1-2] classified as "not relevant".

E. Results

We evaluated 42 items among the three proposed interfaces. During the software evaluation we collected information on 46 movies in different contexts, having a data set sufficiently sparse, given that only 4 out of 46 items were not evaluates (9.52% of the data set).

The 46 movies were collected in different contexts. 21 of them (46%) were collected in one movie theatre (Cinemark Alfa) located at Shopping Mall Alfa and 25 (54%) at another movie theatre (Cinemark Beta) located at Shopping Mall Beta, both located at the same city (OurCity-AnyState), s shown in figure 9.

Figura 10 illustrates the frequency the users visited each shopping mall.

Out of the 71 movies that were showing at all the Cinemark network during the experiment period, only 46 were available at the collected contexts, that is, FBIC filtered 25 items that were not available at the collected contexts (100% efficiency), reducing the sample by 35% because those items were not adequate to the user context obtaining and efficiency that was 14% higher for context information (in case FBIC was not applied, only 75% of the available items at that time period would be adequate to the context).

Figure 11 illustrates the amount of items evaluated per technique de técnica (FBP, FC, FBIC) and sex, as available in the prototype, using the collected sample. At this graph we can
see that the female public tends to use more the recommendation the prototype offers, but this affirmation is not conclusive because of the small sample. Nevertheless, this may be an indication that the female public may be more prone to use the functionalities of this kind of application, even though this need to be confirmed by further studies.

At the evaluation process, we evaluated three algorithms: FC (the classic collaborative filtering technique), FBIC and FBP.

After the users performed all the experiment steps, its data were processed and we calculated the MAE and precision of both algorithms, as we can see in Figure 12.

The collected values corroborate the idea that the system using FBP performed better than FC and FBIC. MAE was 9% smaller and precision 10% higher using FBP when compared to FC. This variation, presented in figure 12, represents only the difference found in the experiment, given that it was not the goal of this work to detect any superiority of any given method, but our data can be used for more detailed posterior analysis.

Fig. 10. Frequency of visit of the sample to movie theatres

The collected values corroborate the idea that the system using FBP performed better than FC and FBIC. MAE was 9% smaller and precision 10% higher using FBP when compared to FC. This variation, presented in figure 12, represents only the difference found in the experiment, given that it was not the goal of this work to detect any superiority of any given method, but our data can be used for more detailed posterior analysis.

Fig. 11. Item evaluation by type of technique

F. Discussion

Even though it was not the goal of this work to prove the superior quality of the recommendations generated, it was possible to realize that Personal_Movie applying FBP in the recommendation had a precision that was 10% higher and MAE that was 9% smaller than FC (the traditional approach) in the user sample selected. This different can be attributed to the use of FBIC that is proven to increase other techniques efficiency when applied together, generating better recommendations. Through the application of FBIC, it was possible to remove 35% of the items that were not suitable to the user contexts.

The data presented show an improvement in item recommendation through FBIC application together with the techniques approached in this work (FC and FBP), but this work cannot be taken as a basis for proving this kind of superiority. Nevertheless, the data is available as a possible subsidy for posterior analysis.

VI. CONCLUSIONS

In the next years, Brazil will host the World Cup, the Olympic Games and the Paralympic Games. During those events that is a bigger demand for personalized information on the attractions offered by the host country, such as movies. Given this context, the prototype is relevant and has the potential to take personalized information to the tourists in loco.

Giving personalized information to the tourists allows us to hope they will be pleased with the offered services, increasing the probability of them using the national product or service, heating up the economy, as well as providing subsidies for them to form a positive image of the country and the desire to come back for another visit.

A. Future Work

Given the results found, it was possible to verify the potential of Personal_Movie to provide efficiently personalized information to its target audience. As a future work we intend to deal with the issues that we believe may also be important during megaevents such as (i) application of techniques for group recommendations; (ii) creation of a personalized graphic interface for the user based on his personality traits; (iii) availability of the software in other
mobile platforms; (iv) mapping of personality traits implicitly through the evaluation of the movies the user has already watched; (v) analysis of the requisites necessary for system scalability (vi) verification of the data in bigger populations in order to better understand the statistical significance of the results found.

VII. REFERENCES


