

# Content Personalization System Based on User Profiling for Mobile Broadcasting Television

Silvia Uribe<sup>1</sup>, Iago Fernández-Cedrón<sup>1</sup>, Federico Álvarez<sup>1</sup>,  
José Manuel Menéndez<sup>1</sup>, and José Luis Núñez<sup>2</sup>

<sup>1</sup> E.T.S.I. Telecomunicación, U.P.M., Avda Complutense s/n, 28040 Madrid, Spain

<sup>2</sup> Zentym Solutions, Madrid, Spain

{sum, iff, fag, jmm}@gatv.ssr.upm.es,  
jln@zentym.com

**Abstract.** Content personalization is a key element in the media content environment since it contributes to improve the user's experience. In this paper we present a novel intelligent content personalization system for content flow personalization over mobile broadcasting networks and terminals based on user profiling and clustering taking advantage of the consumption data obtained from the user and the information given by user's tastes and behavior.

**Keywords:** content personalization, mobile television, clustering, profile segmentation, preference, DVB-H, interactivity.

## 1 Introduction

The digitalization of the media environment in broadcasting networks has opened new possibilities and at the same time, it presents new challenges to the actors involved in the value chain. In fact, consumers expect new incentives in order to contribute to this progress, and for this reason it is needed the development of innovative services that help them to satisfy their needs and to fulfill their expectations.

In this paper we present a novel intelligent content personalization system for content flow personalization over mobile broadcasting networks and terminals based on user profiling and clustering taking advantage of the consumption data obtained from the user and the information given by user's tastes and behavior.

One of the most common ways to improve this experience is by developing content recommendation systems based on user's preferences and behavior. In fact, there are many techniques to implement these recommendations engines: user clustering and profile classification, which are not only used in media environment but also in the web [1] or television applications [2]. But the main difference between these applications and the intelligent system for content personalization presented in this paper is that our system provides a final, clear and transparent solution for the users, by allowing them to receive the content according to their profiles, and not making necessary them to choose the piece of content to be displayed. In this way, since it is developed for mobile broadcasting television, they do not have to choose between several offered options, thus obtaining a more efficient system and real time results.

The rest of the paper is organized as follows. Section 2 outlines the state of the art of recommendation and personalization systems and the techniques used on them.

Next section shows the users' clustering design for the solution in this paper. Section 4 presents the general structure of the system and the different modules. The system operation is shown on section 5 and in section 6 we present the conclusions.

## 2 Overview of the Prior Art in Recommendation and Personalization Systems in Media Environments

One of the most efficient solutions to this problem is the so called recommendation system. In fact, as it was explained in [3], the appearance of new television standards such as DVB-T or DVB-S increased the number of programmes and channels and one effective solution is helping them to find out their favorite contents by giving them several recommendation based on their preferences and tastes. According to this work, these systems are typically composed of four modules:

- User profiling module, which is in charge of generate the profiles of users according to the existing information.
- Program modeling module, which is in charge of extract the content data to classify them.
- Collaborative filtering/content based module: this module creates user's groups by finding out user's neighbors with similar preferences or behaviors or chooses the content to recommend by studying its characteristics and their similarity to user profile. In [4] authors show a wide classification of the existing recommendation techniques.
- Recommendation module: the final module which match content and users.

Within recommendation systems, one of the most important research areas is the content and user activity modeling, since there are many papers focused on the content modeling to allow the recommendation process, mainly based on semantic process of content metadata,, as it is explained in [5] and [6].

But content recommendation systems are also applied not only to television, but also to other environments like web browsing related to media content. In [7], authors present a system that helps users to find out media contents while browsing the Internet. In this case, the content recommended lists are created by analyzing content metadata and users' feedback.

As we have just seen, content recommendation systems are innovative solutions to integrate users' needs and preferences in the media chain. But these solutions require the users to finally make the decision of which content is going to be selected, or in other words, they are non transparent solutions for the users because they need an action by the user. For this reason, the development of content personalization system represents an evolution which contributes to obtain more efficient solutions for real time applications.

Moreover, personalization system can focus on different parts. As it is explained in [8], a structure personalization is needed in order to allow users to access the same content in different devices with different capabilities such as PC, mobile devices and so on. But as well as the structure personalization, content personalization itself represent other way of providing personal flow to the user. In [9], authors focus on the personalization of interactive video content in sports events, but the system presents in this paper does not only focus in a specific kind of events, and it is initially designed to be applied in mobile television, although it could be extended to other environments like other digital television standards and IPTV.

### 3 Design of a Content Personalization System for Mobile Television: The User Profiling and Content Association

The aim of this system is to improve the user experience by providing targeted content to users according to their preferences, in an automated and real-time way.

User preferences, that is, the center of this system, can be obtained by many ways, as it is explained in [10], where they are predicted by considering user's conformity, user's context and finally user's behavior.

In our system, according to both the application architecture and the available information, the most efficient and effective way to obtain the user's clustering is by the generation of different users' profiles, according to the declared preferences and to the media consumption, and cluster them according to pre-assigned groups.

Given the characteristics of the system, and in order to let the designer of the system to establish threshold values for the belongingness to a cluster of a user we use the fuzzy c-means algorithm. Besides, this solution makes the matching between users and content easier, because advertiser and content provider generate concrete users' classes with specific associated contents.

In this case, the number of groups or clusters (with their centroids) is set a priori.

The initialization of the clustering method is composed of 5 steps, according to the final desired classification:

1. First of all, users have to complete a set of questions about their habits, interests and preferences. Each question is designed to have four different answers, each of them with a different score.
2. Then, we collect the information about media consumption to elaborate a user profile by mixing the result from the declared interests and the result from the measured consumption.
3. We apply the clustering algorithm to obtain the centroids.
4. Once these centroids are established, we keep them fixed and we assign the users to each of the centroids.
5. If a new user joins the system, he/she will be assigned to each of the clusters. When the number of new users reaches the update threshold (UT) then we will perform again the step number 3.

Mathematically, the process from 1 to 5 is explained as follows:

In the first step, 1, we define a set of users  $U = \{U_1, U_2, \dots, U_n\}$ , holding a specific profile  $U_p$ . Then we create the function to model the interests and preferences of a user  $U$  according to the number of content categories  $N$ :

$$U_{pi} = \sum_{n=1}^N L(U_x) | C_n \quad (1)$$

Where  $L(U)$  is the declared interest in a scale of 0 to 1.

Then we can model the media consumed ( $M$ ) according to the total time consumed ( $T$ ) per content category ( $C_n$ ) and the number of users  $U$ . The result is adapted to a scale from 0 to 1, to be compared with the  $U_{pi}$

$$U_{Si} = \sum_{n=1}^N M_A(U_i | (T, C_n)) \quad (2)$$

In order to mix both measures, it is needed to establish a weight for each of the components, which can be established by the designer of the system. Then, we compute the fuzzy c-means algorithm, to obtain the centroids and the membership values of the different users to assign them to a cluster. The clustering (in the step 3) is computed through the minimization of the objective function:

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2, 1 \leq m \leq \infty \quad (3)$$

Where:

- $m > 1, m \in \mathfrak{R}$  ;
- $u_{ij}$  is the degree of membership of  $x_i$  in the cluster  $j$ ;
- $x_i$  the  $i$ -th term of the n-dimensional input data;
- $c_j$  the n-dimension centre of the cluster
- and  $\|\cdot\|$  is any norm measuring the similarity between the input data and the centre.

The process is iterative and the optimization of the objective function  $J$  is done updating the membership ( $u_{ij}$ ) and the cluster centres ( $c_j$ ) as follows:

$$u_{ij} = \frac{1}{\sum_{k=1}^C \left( \frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}} \quad (4)$$

$$c_j = \frac{\sum_{i=1}^N u_{ij}^m x_i}{\sum_{i=1}^N u_{ij}^m} \quad (5)$$

The iteration stops when  $\max_{ij} \left\{ |u_{ij}^{(k+1)} - u_{ij}^{(k)}| \right\} < \varepsilon$  where  $\varepsilon$  is a termination criteria between 0 and 1, with iteration step  $k$ .

According to the system structure, content providers use the server module on the server side not only for uploading new content, but also for associating each one of these contents to the different existing users' classes. Thanks to this solution, there is no need of an automatic association or even a content modeling process in order to extract content characteristics, making the system operation simpler.

## 4 Content Personalization over Mobile Television: System Architecture

A content personalization system over mobile television does not only involve the final users and their preferences, but also the other actors in the media value chain. As it is shown in Fig. 1, in this system there are many components that have to work together in order to obtain the best and most efficient result.

The proposed system is divided into three different parts according to their functionalities: first of all, the broadcaster, which is in charge of providing the content and giving the timestamps of the next change in order to make the swap between common content and personalized one at the right time; second, the personalized content server, which is in charge of doing the assignment and management of the personal content, based on the user clustering and the user-content assignment; and finally, the user side, the user device, which must present the given information and it also allows the bidirectional communication via return channel.

Next subsections explain the functionalities of each architecture module (Fig. 2).

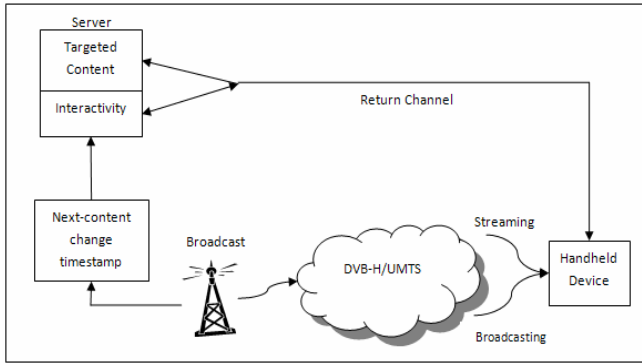


Fig. 1. System architecture

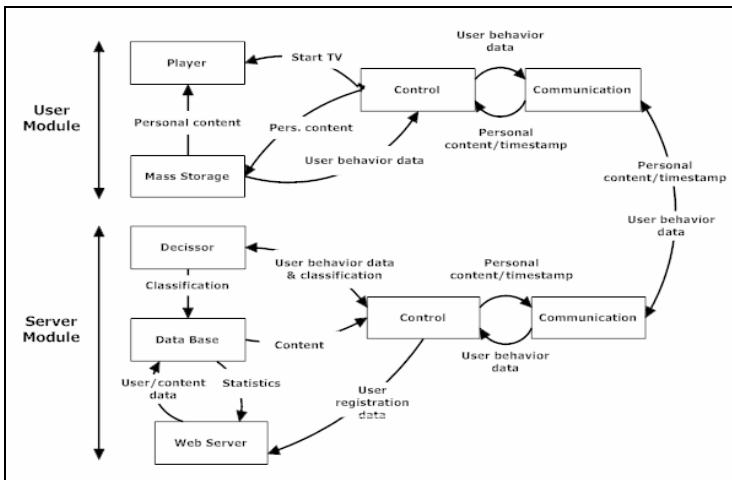


Fig. 2. System modules' interaction

## 4.1 Broadcaster

Broadcasters are in charge of providing the content flow. In order to replace the common content with the personal one at the right time, a timestamp of commercial breaks is needed. Although there are different algorithms to detect spots in media flows, like the one shown in [12], we consider that these solutions present a high process load for mobile devices, and that is why we propose the broadcaster to include this information, which can be done in two different ways:

- On one side, broadcasters can report to the users the time left to the next break via the return channel. This is the implemented solution.
- On other side, this information can be included in the ESG (Electronic Service Guide), but this solution requires both a real-time update of the guide as well as a constant scan of it in the user side in order to notice the changes. This method is processor intensive and consumes a huge amount of battery, so at the moment is not recommended for a real time application.

## 4.2 User Module

This module has two main objectives: the content presentation and the implementation of a bidirectional communication channel, and it is divided into different modules:

- Player: this module presents the broadcasted content and the personalized one.
- Mass Storage module: it stores both the personalized content received from the server and the user's behavior data, which can be used to update users' profiles.
- Control module: it is in charge of presenting user's forms for the personalization and registration processes. Once users are logged in, this module starts the TV presentation, asks for the next commercial break and captures users' interactions.
- Communications module: it is in charge of receiving the personal content, and sending back the collected data to the server via an UDP socket.

## 4.3 Server Module

The main function of this module is making the user clustering and segmentation in order to relate each user's group to the content. This module is divided into the following modules:

- Web server: in this server broadcasters and content providers have the tools to manage the system, by controlling the users' profiles (generation and update) according to the existing information, by creating new commercial campaigns and specifying the target group for them. In this server, broadcasters also find the users' consumptions statistics.
- Media Asset Decision module: based on the method explained before, this module associates each user to the correct group and then selects the tailored content for them. This module has the control of the personalization, and can update the users' profiles.
- Communications module: this module controls the data flow between the server and the user.

- Control module: this module allows the data flow inside the server by giving it the proper format.
- Database: it stores the personal profiles, the data collected from user's behavior, as well as the content given by the content providers.

## 5 Content Personalization over Mobile Television: System Operation

Based on the structure of the system, its operation (shown of Fig. 3), can be divided into six steps:

1. Clients log in the system, by sending users' connection data (user&password), in order to be identified in the system. Once the server cross-check this information on the database, the user is connected and starts watching TV. If it is the first time of the user in the system and he has completed the first registration form to minimize the so called 'cold start' effect, server receives these data and performs the user's segmentation according to the clustering method designed.
2. Once connected, client device asks the server for both the exact time of the next commercial break, and the personalized content to show.
3. The actions performed by the users are monitored in order to update user's profile. These data are stored on the client side until they are sent to the server side via the return channel.

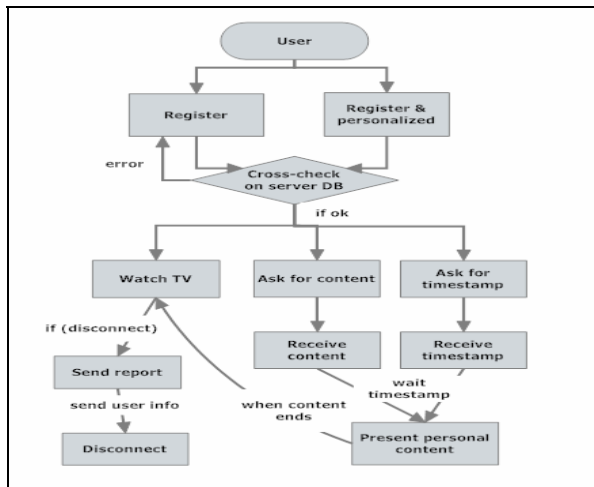


Fig. 3. System operation

4. Once this information is received, it is processed to update the user's personal profile and perform the most efficient personalization.
5. According to the exact time received about the next commercial break, the presentation module on the user side swaps between the common content and

the previous stored personalized one, without any action by the user. Later, when the personalization ends, the system returns to the common media flow.

6. Finally, in order to log out the system, users have to send again their connection information.

## 6 Conclusions and Future Work

We have presented an efficient system for content personalization. It is based on user clustering according to their preferences and behavior, which means that user's environments are the center of the development. Furthermore, this system represents a useful solution for costumer segmentation and personal advertising, which can be applied in other environments. As it has been explained, user's segmentation includes two main steps: the first one where each user is classified in a group according to his answers to the proposed set of questions, and the second using their content consumption historic data. The system has been extensively tested and is in prototyping by the time of this paper submission.

## References

1. Phatak, D., Mulvaney, R.: Clustering for Personalized Mobile Web Usage. In: International Conference on Fuzzy Systems (2002)
2. Virvou, M., Savvopoulos, A.: An intelligent TV shopping Application that provides Recommendation. In: 19th International Conference on Tools with Artificial Intelligence (2007)
3. Xu, J., Zhang, L.-J., Lu, H., Li, Y.: The development and prospect of personalized TV program recommendation system. In: IEEE Fourth International Symposium on Multimedia Software Engineering (2002)
4. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering* 17(6), 734–749 (2005)
5. Bellekens, P., Houben, G., Aroyo, L., Schaap, K., Kaptein, A.: User model elicitation and enrichment for context- sensitive personalization in a multiplatform TV environment. In: 7th European Conference on Interactive Television (2009)
6. Velusamy, S., Gopal, L., Bhatnagar, S., Varadarajan, S.: An efficient ad recommendation system for TV programs. *Multimedia System Journal* 14(2), 73–87 (2008)
7. Liu, Y., Yang, Z., Deng, X., Bu, J., Chen, C.: Media Browsing for Mobile Devices based on Resolution Adaptative Recommendation. In: International Conference on Communications and Mobile Computing (2009)
8. Yin, X., Lee, W.S., Tan, Z.: Personalization of web content for wireless mobile device. In: Wireless Communication & Networking Conference (2004)
9. Mylonas, P., Karpouzis, K., Andreou, G., Kollias, S.: Towards an integrated personalized interactive video environment. In: IEEE Sixth International Symposium on Multimedia Software Engineering (2004)
10. Boutemedjet, S., Ziou, D.: A graphical Model for Context-Aware Visual Content Recommendation. *IEEE Transactions on Multimedia* 10(1) (January 2008)
11. Jain, A.K., Murty, M.N.: Data clustering, a review. *ACM Computing Surveys* (1999)
12. Covell, M., Baluja, S., Fink, M.: Detecting Ads in Video Streams Using Acoustic and Visual Cues. *Computer* 39(12), 135–137 (2006)