

Analysis of different Approaches of Targeted Advertisement

Abhijeet Shingade
Department of Computer
Engineering
Sinhgad Academy of
Engineering

Adesh Atole
Department of Computer
Engineering
Sinhgad Academy of
Engineering

Piyush Galphat
Department of Computer
Engineering
Sinhgad Academy of
Engineering

Shashank Dharmadhikari
Department of Computer Engineering
Sinhgad Academy of Engineering

Bhushan Thakare
Department of Computer Engineering
Sinhgad Academy of Engineering

ABSTRACT

Many businesses have advertisement as their primary source of revenue. Print and Audio Visual media are on forefront in this area. In today's technology obsessed world, advertisement also forms a major source of revenue for service providers, app developers who work on mobile platforms. Advertisement on mobile platforms differs a lot from traditional advertisement. The information available with mobile devices can be exploited to target the advertisements towards specific audience. This paper discusses various approaches for targeted advertisements. It also discusses an approach to optimize allocation of advertisements to target audience.

Keywords

advertiser, Bayesian network, mobile devices, publisher, targeted advertisement.

1. INTRODUCTION

Advertising is a major source of revenue for many services such as news media, entertainment industry and television. There are many techniques through which personalized advertisements can be shown to a specific group of audience. Many traditional approaches focus on targeting a group of people, such as print media advertisements. The publishers assume the properties of the audience, who will be accessing their content and based on the assumed type of the audience, they publish the advertisements. But, the problem in this approach is the assumptions are not always right, the advertised content may not reach to the intended audience expected by the advertiser.

There is a lot of difference between advertising on mobile devices and traditional advertisements in print media and advertisements on television. A lot of user specific and environment specific information is available with mobile devices. Advertisement publishers and advertisement publishers can exploit this data in order to show right advertisement to the right person in right context and at the right time and location.

2. MOBILE TV TARGETED ADVERTISEMENT AND CONTENT PERSONALIZATION

The digital multimedia technology has evolved so much from its birth to present time it has gone through revolutionary changes, handheld devices have a significant transformation from a device that provided voice service to complex devices with multiple functionalities.

2.1 Need of new service model for advertisement

The re-broadcasting of ordinary TV has the disadvantages of smaller display size, low battery and processing power of mobile devices, so to overcome this disadvantage a new model is tailored where user receive tailored contents adapted to their User Profile.

2.2 Basic dimensions of Personalization

1. Structure personalization: mainly focuses on the content configuration like size of content, font used for content, etc. this personalization basically focuses on the users' choice of viewing the content.
2. Content personalization: focuses on user preference, client model uses, etc.

2.3 Personalization Levels

1. User side: personalization is done on user's side i.e. on the user's device.
2. Server side: user's preferences, content consumptions, etc.
3. Both side: by combining the advantages of both the models.

2.4 System Architecture

The first part is the module in charge of providing the regular broadcasting, this can be one by available broadcasting network such as DVB-H, or by multimedia streaming over a mobile communications network such as UMTS. The way of content delivery also depends on the capabilities of the users devices. The server where the television flow and the data services are located, is also considered as a part of the broadcasting module [1].

1. The second part of the whole architecture is the personalized content server. This server can also be used to provide interactivity between system and users, via return channel, which is mostly implemented over available mobile network. The content server is in charge of doing the assignment and management of the personal content, according to UP based on their preferences and their consumption data, which can be obtained via return channel. It also indicates when the change between common and personalized content has to be made by sending to the client part the time left for it [1].
2. Personal content is sent over mobile network, since it has to get to each user separately. So, in short, this 'unicast path' via mobile network is used for personal content supply as well as a back channel interactivity which cannot be provided by the broadcast path [1].
3. Finally, user part can be found out, consisting of the user's devices that can both execute the application and send back the needed user data that help configuring their user profile [1].

2.5 System Modules

The System architecture shown in Fig. 1 is divided in two parts on basis of their functionalities and capabilities as shown in Fig. 2. The two parts are user side and server side.

1. User side has the following modules:
 - a. Player: it takes care of presenting the broadcasted content and personalized content.
 - b. Mass Storage module (DB): it stores the content received from server via mobile network and the collected data about user behavior.
 - c. Control module: it controls the presentation and it monitors and captures user's interaction and prepares these data to send via the communication module.
 - d. Communication module: it receives content from the server and sends the user behavior back to server.

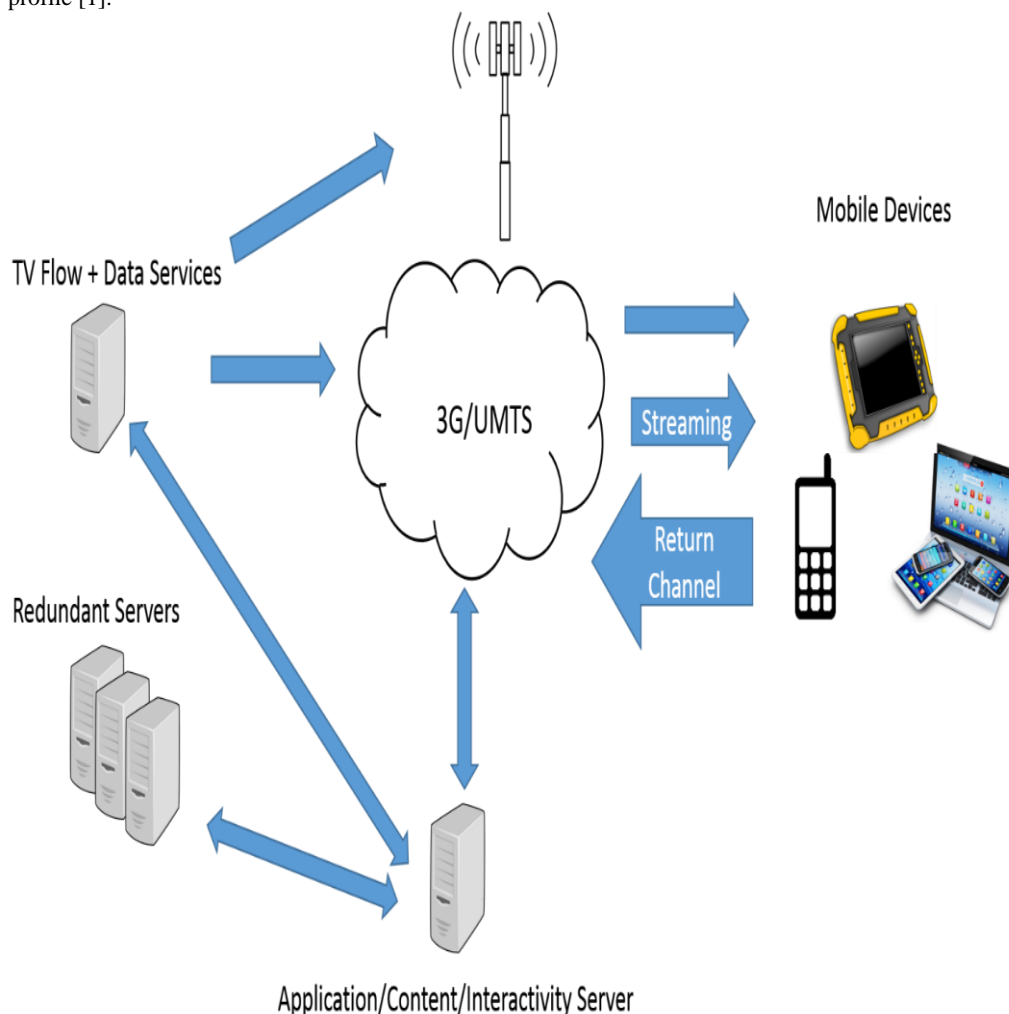


Fig 1 System Architecture and functionalities

2. Server side modules:
 - a. Decissor: it controls the users' profiles on the database module, it updates the user profile information.
 - b. Web server: hosts the web services to manage the system and contents on it.

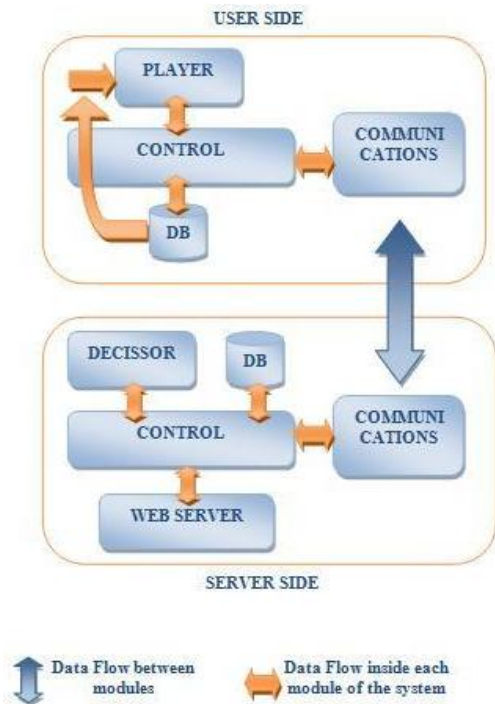


Fig 2 Personalization System Modules

- c. Communications module: this module controls the data flow between the server and the user.
- d. Control module: controls the data flow inside the server. It gives format to the data so modules could communicate safely and in proper way.
- e. Database: stores the personal profiles and the data collected from users' behavior.

3. PUSH ADVERTISEMENT ON MOBILE DEVICES

Advertisement can be broadly divided into two types:

1. Push Type Advertisement Push type of advertisement forcibly brings content to the user. E.g. Email blasts, TV, Radio, Telemarketing etc.
2. Pull Type Advertisement A Goal of Pull type of advertisement is to get consumers come to you. E.g. Social Media, Blogging, Search Marketing, EBooks etc.

This approach trying to overcome present drawbacks in mobile advertising system. At present, mobile advertisement system lacks in publishing right advertisement to right people at right location. After analyzing this, Mohammad Imran M & G Kavitha [2] propose an efficient and high scalable digital advertisement system on cloud for location and user profile based mobile advertising [2]. This approach introduces new idea of using the advertisement data on a cloud connect with mobile user via different mobile devices.

3.1 Architectural Framework

Architecture consists of four modules

1. Obtaining User and location
2. Ad's Selection & Publishing
3. Updating Click Streams & User Behavior
4. Performing analytics on a data

1. Obtaining User and Location

User and location information is obtained in this module of an architecture. User profile information is obtained from social media like Facebook, twitter, LinkedIn etc. The profile information of the user such as name, age, gender, mail id used by system if user set these information available for public in the social media settings.

The location co-ordinates obtains from longitude and latitude co-ordinates of mobile base station under which user mobile node is connected. Location information is changed when user transfers from one cell to another cell.

2. Ad's selection and publishing

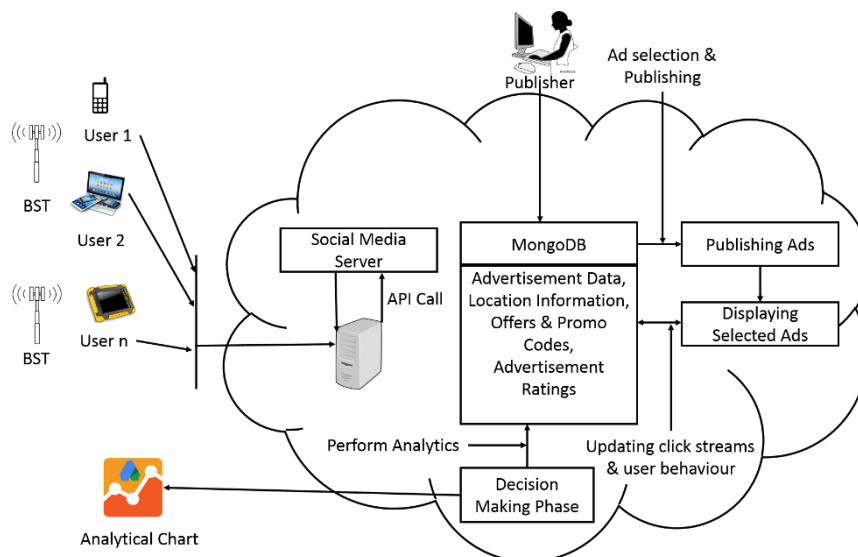


Fig 3 Push Advertisement Architecture

In this approach, geofencing technique is used. In geofencing technique the obtained location co-ordinate is used as a center. The search query search

for the advertisement data only within the specified radius of the mobile base station location

coordinates. The result of these data are display in the space of advertisement.

$$\Sigma^* = \sum (PI_{i,j} || LI_{[lon,lat]}, (LI_{[lon,lat]}, r))$$

Σ = Search query

Σ^* = Search Result

PI = Profile Information

Where, i represents id, j represents gender, age, mail id etc.

LI= Location Information

Where, lon and lat represents longitude and latitude respectively.

3. Updating click streams and User behavior

In this module click-streams are and in-depth clicking are analyzed. In depth clicking is analyses as follows:-

Let us consider, that advertisement has "N" pages and "PIij(Ck)" represent particular user "PIij" and "C" represents the click of user and "k" represent number of clicks. Say k=1, 2, 3, 4 . If PIij(Ck) == N represent that the particular user have reached the end of the page and assumes that he might be interested to buy the product.

Let us consider that advertisement has three pages and initial display starts with zero

Say,

N = 3,

$C_k = C_0$, where k=0, 1, 2, ..

If user made a click on the advertisement initially

Then,

$C_k = C_k + 1$

$C_k = C_1$

Let us consider corresponding description to every click value made by the user which can be used for the analytics purpose.

Let say,

K = {"initial display", "full view mode", "purchase"}

Now, $C_k=1$ or C_1 then it defines that that the user have made only one click which reaches the first page. It can be represented as PIij($C_k=1$).

4. Performing Analytics on data

Here users data such as click streams, user ratings etc. used to perform an analysis. From this analysis, advertiser can change their advertisement pattern. In this approach user profile based analytics also carried out. Depending upon average user age and a time user has mostly visit advertisement, they made a graphical representations for an easy understandings of advertisement publisher.

4. LOCATION BASED HYPER LOCAL TARGETED ADVERTISEMENT ON MOBILE DEVICES

With a recent boom in number of smartphone users, advertising agencies are adapting themselves to provide targeted advertising solutions to advertisers. Many content developers are also exploiting this to their advantage, which

leads to higher number of impressions/clicks leading to increased revenue for all stakeholders. But this form of targeted advertisement is still not penetrated in local markets, where local businesses are willing to exploit the advantages of targeted advertising. A hyperlocal approach aims to focus in this area and making advertisements more helpful to viewers than annoying. Location information of device plays the most vital role in this approach.

Following are the pieces of information, which will play vital role in hyperlocal targeted advertisement:

1. User's personal information
2. Location
3. User's daily usage patterns
4. Metadata about the content developed by publisher
5. Metadata about the advertisement given by publisher

4.1 Targeted Adaptive Advertisement

The advertisements will be shown to the end users, which fully or partially comply with the requirements specified by the publisher. Thus, a publisher can choose the intended audience in case of targeted advertising, which is not possible with traditional form of advertising, if multiple advertisements satisfy with the user's context, ads can be scheduled using a queue.

An essential form of advertising known as Targeted Adaptive Advertisement (TAAD) as discussed in Mazaheri [3] can be used to decide the audience for particular advertisement.

4.2 Architecture

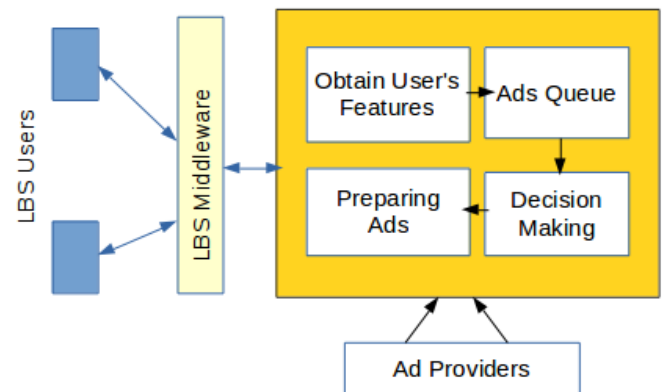


Figure 4 TAAD Architecture

The image shows a typical architecture of TAAD [3]

Bayesian networks can be used to extract user's features. The central component of TAAD is Decision Making, which ranks the advertisement based on user features. The 'Prepare Advertisement' component simply fetches the ad, which has highest rank.

The user features can be extracted using Bayesian networks. This approach for Bayesian networks is discussed in detail in Mazaheri, Rafiee, Khadivi [3] and Pearl [4]

The ranking of advertisements can be done using Fuzzy TOPSIS as discussed in Mazaheri, Rafiee, and Khadivi [3]

5. USE OF BAYESIAN NETWORKS FOR OPTIMIZING ALLOCATION OF ADVERTISEMENTS

A Bayesian network is a probabilistic graphical model that represents a set of variables and probabilistic independencies. It is a probabilistic model that is used to represent uncertainty knowledge and show the causal relationship among variables [5][6]. Recently many researchers are trying to develop Bayesian network models for learning from data. For example, Heckerman provided a tutorial on Bayesian networks and associated Bayesian techniques for extracting and encoding knowledge from data [7]. In this case, Bayesian network model can be used to allocate right advertisement when a web user visits the page based on a large amount of historical web session data.

5.1 Data Availability And Acquisition For Bayesian Network Model Implementation

The Bayesian model needs data to process for that purpose extracting user data through web logs, cookies and user profiles is being done. However without users' permission their web logs can not be accessed so in that case cookies are the best way to extract users' data. Cookies are parcels of text sent by a server to a web browser and then sent back unchanged by the browser each time it accesses that server. They become accessible to the same server if a user browses the same website again. Cookies have been widely used for authenticating, tracking, and maintaining specific information about users, such as site preferences and the contents of their electronic shopping carts. The third type of data source is user profiles. Although user profiles are less available because many visitors of a website are not registered users, once available they are informative in improving the accuracy of targeted advertising.

5.2 Bayesian network model for advertisements allocation [8]

In this all types of advertisement are classified into clusters. When the web log document or cookies information are retrieved, the useful information which is valuable to us to build and train classification models must be extracted. The text format data such as web log and cookies generally contain some fields, such as session ID (the client address or url), access date and time, the url of the accessed page, etc., which are in the sequential order. This information allows us to derive the attributes such as user's address, session start time, average page, average page time, and login total second and so on. Initially construct a Bayesian network model and determine some of the possible variables to model the study.

In Bayesian network model, let C represent the web user click the advertisement or not, $C = \{c1 \text{ "Yes"}, c2 \text{ "No"}\}$. $C=c1$ means the web user clicks the advertisement, and $C=c2$ means the web user does not click the advertisement. Use ST to denote the time when the session starts. $ST = \{st1 \text{ "Morning"}, st2 \text{ "Afternoon"}, st3 \text{ "Night"}, st4 \text{ "Midnight"}\}$. Therefore, ST can be "Morning" when start time is from 6 am to 12am, "Afternoon" when start time is from 12 am to 6 pm, "Night" when start time is from 6 pm to 12 pm or "Midnight" when start time is from 12 pm to 6 am. Similarly, use AT to denote the average time the web user spent for a page. $AT = \{at1 \text{ "Long"}, at2 \text{ "Short"}\}$. P denotes total number of pages the web user visited. $P = \{p1 \text{ "Many"}, p2 \text{ "Few"}\}$. S denotes search or not. $S = \{s1 \text{ "Yes"}, s2 \text{ "No"}\}$. TT is total time of this

session. $TT = \{tt1 \text{ "Long"}, tt2 \text{ "Short"}\}$. The model is shown in Figure 5.

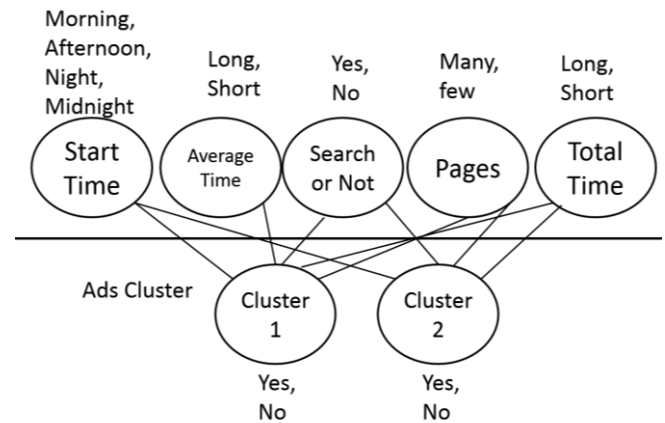


Fig 5 Bayesian Network Model for Targeted Advertisement

The structure and probabilities of a Bayesian network can be learned from data [7]. From historical data, the probabilities of each value in all variables can be found such as $P(ST=st1)$, $P(C=c1)$ and so on. Suppose there are n clusters. For each cluster calculate the conditional probabilities:

$$P_i(C|ST,AT,S,P,TT) = \frac{P(C,ST,AT,S,P,TT)}{P(ST,AT,S,P,TT)}$$

In the model, there are five variables about web users' online behavior and each variable has two or four kinds of values. Therefore, there are 64 ($4*2*2*2*2$) kinds of web users' online behavior. They can called as web user behavior patterns. $P(ST,AT,S,P,TT)$ represents the probability of each pattern. Because ST, AT, S, P, TT are independent, so:

$$P(ST,AT,S,P,TT) = P(ST)P(AT)P(S)P(P)P(TT) \quad (2)$$

(1) and (2) are used to get (3):

$$P_i(C|ST,AT,S,P,TT) = \frac{P(C,ST,AT,S,P,TT)}{P(ST)P(AT)P(S)P(P)P(TT)}$$

In formula (3), all probabilities on the right of equation can be calculated from historical data.

Therefore, with each pattern, a series of conditional probabilities $P_i(c1|ST,AT,S,P,TT)$ can be found, for example $P_i(c1|st1,at1,s1,p1,tt1)$. This is the probability that web users click the advertisements in cluster 1 when their start time of visiting is in the morning, the average time is long, they visit search page, they visit a lot of web pages, and the total time is long. Also there are other 63 probabilities in other conditions. Those probabilities are important data to allocate advertisements. After getting all probabilities of all advertisement clusters, sort all n probabilities of each web user behavior patterns, for example $P_i(c1|st1, at1, s1 ,p1,tt1)$ $i = 1,2,\dots,n$. Then which cluster is the best suitable to which kind of web users can be found out.

6. CONCLUSION

Thus, this paper provides various approaches to implement Targeted advertisement, which is in favor of both advertisers and audience. It allows publishers to target advertisement to specific audience. Mobile devices perfectly accompany targeted advertisement, which benefits advertisers and

publishers. Use of Bayesian Networks allow optimal scheduling of Advertisements based on user's profile.

7. REFERENCES

- [1] Silvia Uribe, Iago Fernández-Cedrón, Federico Álvarez, José Manuel Menéndez, José Luis Núñez, and Mobile TV Targeted Advertisement and Content Personalization, IWSSIP 2009. 16th International Conference on Systems, Signals and Image Processing.
- [2] Data for Targeted Mobile Mohamed Imran M, G. Kavitha, "An Efficient Utilization of Advertisement Advertisement System", 2014 International Conference on Advanced Communication Control and Computing Technologies (ICACCCT).
- [3] A. Mazaheri, N. Rafiee, P. Khadivi, Location based targeted advertising using Bayesian Network and Fuzzy TOPSIS, 2010, 5th International Symposium on Telecommunication.
- [4] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann, 1988.
- [5] Pearl, J. Probabilistic Reasoning in Intelligent Systems. San Francisco: Morgan Kaufmann Publishers, 1988.
- [6] Jensen, F. An Introduction to Bayesian Networks. New York: SpringerVerlag, 1996.
- [7] Heckerman, D. A Tutorial on Learning with Bayesian Networks in Learning in Graphical Models, MIT Press, Cambridge, MA, 1999.
- [8] Kai Li, Jianyuan Yan, Xiaowen Wang, A Bayesian Network Model for Optimizing Advertisements Allocation in Intermediate Online Targeted Advertising, 2008 IEEE.