Abstract

The effectiveness of television advertising is diminishing since 1965. Nowadays, the explosion of available channels, the advent of Video-on-Demand and the proliferation of Personal Video Recorders are further lowering the efficacy of television advertisements. As television advertising gets ignored by viewers, internet advertising achieves new profit records. Internet advertising success is tightly related to its relevance, i.e., internet advertisements are often interesting to users. Therefore, it is mandatory to present interesting and relevant advertisements to television viewers as well.

In order to present a solution to this problem, we have designed, implemented and evaluated TargetAd (Targeted Advertising), a system for targeting advertisements on IPTV systems. TargetAd’s goal is to present advertisements which are relevant to the user. Additionally, it is shown that targeted television advertising generates greater revenues than traditional advertising, improves the effectiveness of advertisers’ investments and, most importantly, betters users’ television experience.

Keywords: IPTV, Targeted Advertising, Television Advertising, Advertisements, User Modeling

1 Introduction

In 1965, 34% of the american television viewers were able to remember a brand which was advertised during the television show they had just saw. In 2000 this percentage has fallen to 9% [3]. Therefore, television advertising is becoming less effective over the years.

There are several reasons for this lack of effectiveness. The explosion of the number of channels available fragments the audience. Consequently, this fragmentation makes reaching a specific target more difficult, as it may be split throughout several channels. The advent of Video-on-Demand (VoD) and the proliferation of Personal Video Recorders (PVR) brought new forms of ad-skipping, and further contribute to lower television advertising efficacy [3, 11].

The success of an advertisement depends heavily on its relevance to the consumer [3]. The soaring profits of internet advertising reveal the importance of the previous observation. According to last year studies, in the first six months of 2008, internet advertising profits in the United States of America rose 15.2% in comparison to the same period of 2007 [5]. More recently, advertising spending on internet has overtaken television expenditure in the United Kingdom for the first time [6]. It is believed that the accountability of online advertising had a key role in this flip over, as advertisers felt the need to control their campaigns more tightly during tough economic times.

The major difference between internet and television advertising is the relevance of the advertisements to the consumer. A television viewer is often presented with irrelevant advertisements, whereas an internet user very rarely finds uninteresting ads. Thus, it is desirable to make television advertising more relevant...
to the viewer, in order to avoid the current decline of its effectiveness.

The goal of this project was the creation of a system based on Internet Protocol Television (IPTV) that could deliver more relevant television advertisements to its users. The system was named **TargetAd** (Targeted Advertising).

TargetAd seeks to prove that targeted television advertising brings several advantages:

1. A rise in television operators’ advertising profits, as they will take an active role in the advertisements’ spots sales process;

2. An increase in the Return On Investment (ROI)\(^1\) of the advertising campaigns, since advertisements will reach more frequently their targets;

3. An improvement in users’ television experience, because presented advertisements will match users’ interests and preferences more often.

In order to deliver targeted advertising, TargetAd builds users’ profiles based on the analysis of the television content they view. This analysis is only possible if we are able to monitor the contents displayed on a television set. In the case of IPTV, this can be done by installing software in the Set-Top-Boxes (STBs).\(^2\)

TargetAd is also based on the premise that channels send metadata associated with the contents they broadcast. This metadata consists typically on text that describes and classifies the associated content. This service is commonly known as EPG (Electronic Program Guide).

2 **Related Work**

In this section, we will first analyse the six criteria that define a system able to deliver targeted television advertising. Following this analysis, we will make an overview of the academic and commercial systems that, somehow, deliver targeted advertising.

The systems that are able to present targeted advertising to their users may be defined using six criteria:

1. **Dynamic Insertion of Ads**: how and where are the targeted advertisements inserted in the data stream;

2. **Ads Scheduling**: in which way are ads scheduled inside a television commercial break;

3. **Targeting**: what kind of information is used to target the ads;

4. **Users’ Profiles**: how are users’ profiles built and which data do they contain;

5. **User Identification**: how does the system know which user is currently watching television;

6. **Reports**: how are television advertising campaigns assessed.

2.1 **Dynamic Insertion of Ads**

Targeted television advertising requires the dynamic insertion of an advertisement into the data stream. The dynamic insertion of ads is done by monitoring points in a channel’s emission that signal commercial

---

\(^1\)Return On Investment is a measure to evaluate the success of an investment.

\(^2\)A Set-Top-Box is an electronic device connected to a television set and an external source of signal. A STB produces images and sounds from the external signal which are then provided to the television set.
breaks intended to targeted advertising [11]. These points must be identified and provided by television channels. The monitoring of such points is implemented on software that may be installed on a STB (client-side) or on a server (server-side). When one of these points is detected in a channel’s emission, the emission’s contents are replaced with targeted advertisings. Thus, the ads may be dynamically inserted locally (STB) or remotely (server). The local insertion of ads is only possible if the STB has storage capabilities.

2.2 Ads Scheduling

For each commercial break it is mandatory that a system for targeted television advertising makes an automatic schedule of the ads, as different users may be presented with different ads, which means different schedules. Although of vital importance, we found no solutions in literature for dynamically schedule advertisments. The only solution we found for scheduling targeted advertisements consists of using a static schedule with several ad spots, in which ads are inserted as needed. This solution has, at least, two drawbacks: the channels must make the schedule available a certain amount of time before the start of the corresponding commercial break, and the use of a static schedule do not account the possibility of having ads with different durations.

2.3 Targeting

A system that tailors television advertising for its users needs to employ criteria in order to target advertisements. These criteria form a kind of advertising targeting. The kinds of advertising targeting most used are:

1. **Thematic targeting**: explores the fact that the audience of a channel becomes more alike as the number of channels available increases [3];

2. **Household targeting**: there is a profile for each household and the advertising targeting is based on these profiles [3];

3. **Stereotypes targeting**: system users are stereotyped and advertisements are targeted based on existing stereotypes [2, 8, 7, 9].

2.4 Users’ Profiles

Some of the targeting methods mentioned above make use of users’ profiles, namely household and stereotypes targeting. Users’ profiles may be built — i.e. created and maintained — through several processes and using different data. Therefore, users’ profiles may vary from system to system on:

1. **Construction processes**: users’ profiles may be created and maintained manually, automatically or through an hybrid approach;

2. **Data sources used**: the data inside a user’s profile may be derived from different data sources, e.g. personal data or contents viewed.

2.5 User Identification

Since television sets are multi-user devices, in a sense that they are used simultaneously by several people, it is obligatory to identify who is currently watching television in order to present targeted advertising. This identification may be done implicitly or explicitly.
The explicit identification of users may be achieved by asking periodically which users are currently watching television [2, 8]. On the other hand, the implicit identification of users may be accomplished in two different ways: by analysing the time periods users watch television [4] or by analysing the channels watched [10].

2.6 Reports

A system that enables targeted television advertising, generally possesses the means to produce reports with data such as:

1. To whom an advertisement was presented;
2. What was the percentage of the advertisement that the user actually watched;
3. When was the advertisement presented;
4. What television contents were also watched by this user.

Though the production of these kind of reports is dispensable in such a system, it may reveal quite useful to evaluate advertising campaigns and to prove system’s effectiveness. Therefore, it is understandable that the production of reports has such an important role in many systems for television advertising targeting [3, 2, 8].

2.7 Existing Systems

In the academic field we only found two systems that are able to target television advertising: iMedia [2, 8] and an unnamed system, that we will call it System T for ease of reading [11].

In the commercial solutions field, albeit several systems attempt or claim to deliver targeted television advertisements, there is only one product that genuinely deserves this title: Invidi’s Advatar™. Advatar™ uses demographic data as one of the criteria for targeting advertisements. In fact, Advatar’s users are classified according to four criteria: gender, age, income and geographic location. A small analysis and a comparison between these three systems is presented in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>iMedia</th>
<th>System T</th>
<th>Advatar™</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insertion of Ads</td>
<td>Local (STB)</td>
<td>Local (STB)</td>
<td>Local (STB)</td>
</tr>
<tr>
<td>Ads Scheduling</td>
<td>Centralized and static</td>
<td>Undefined</td>
<td>Undefined</td>
</tr>
<tr>
<td>Targeting</td>
<td>Stereotypes</td>
<td>Behavioral</td>
<td>Stereotypes</td>
</tr>
<tr>
<td>Users’ Profiles</td>
<td>Automatic</td>
<td>Automatic</td>
<td>Automatic</td>
</tr>
<tr>
<td>User Identification</td>
<td>Explicit</td>
<td>Implicit</td>
<td>Undefined</td>
</tr>
<tr>
<td>Reports</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 1: Comparison between iMedia, System T and Advatar™

3 Architecture

TargetAd is a distributed system, with components installed on the client-side (STB) and on the server-side. TargetAd’s architecture is depicted in Figure 1.
TargetAd’s architecture was designed based on the architectural solution IP Multimedia Subsystem (IMS). IMS is seen as a promising solution for easing the development of multimedia services on IP networks, and also for supporting interoperability and convergence of different kinds of networks [1]. The integration of IMS in TargetAd’s architecture had two main goals: to bring TargetAd as close as possible to a commercial product, that could be set to production with minimum architectural changes; and to enable the system with technology that would allow future extensions. IMS is embedded in TargetAd through three modules: Control Sequence Call Function (CSCF), Home Subscriber Server (HSS) and Domain Name Server (DNS).

TargetAd’s architecture and features will be detailed next.

### 3.1 Advertisements Loading, Scheduling and Insertion

TargetAd allows advertisers to submit their ads directly to the system. This can be done using the module TargetAd Loader Client, which is provided to advertisers. As this submission is transmitted over the internet, it is mandatory to digitally sign the messages, in order to prove their authenticity, integrity and non-repudiation. Thus, it is needed to create a pair of cryptographic asymmetric keys for each advertiser. The advertiser is given the private key, which he uses to sign the messages, whereas the public key remains in TargetAd’s infrastructure.

The advertisements submitted by advertisers, once inside TargetAd, will be downloaded by client STBs. The download may be delayed if there is no bandwidth available or no free space on STB’s storage device.

---

http://www.invidi.com/
Commercial breaks intended to targeted advertising are signaled with User Datagram Protocol (UDP) packets. These packets must be sent every second from the beginning of a break until its end. Television channels are responsible for sending these packets, providing in their payload a number that uniquely identifies the corresponding break and the seconds left until its end. The software that monitors these packets is installed on client STBs.

Once a commercial break intended to targeted advertising is successfully identified, the process of ads insertion begins. In order to present targeted advertisements the system needs to know who is currently watching television. The process for user identification is explained further below. The identified users' profiles are used to find the ads that are eligible to be presented in the current break, according to their genders and ages. The selection of ads that will be actually presented in a commercial break is done through an analogy with the Knapsack Problem. In this analogy, the break duration is the knapsack maximum weight, the items that can be put inside the knapsack are the advertisements, such as the items' weights are the advertisements' durations and the items' values correspond to the number of seconds since the advertisement was last displayed. In this way, we try to minimize the difference between a break's duration and the sum of the displayed ads' durations. The items' values were defined as the seconds since the corresponding advertisement was last displayed in order to assure that all ads will eventually be presented.

3.2 Targeting

TargetAd targets its advertisements according to stereotypes. Advertisers are able to set one or more stereotypes as their ads' targets. These stereotypes are defined based on two features: gender and age. Clients are asked to provide these data when they first use their STB, as it is explained next.

3.3 Users' Profiles

As mentioned above, users' profiles are built when a client uses its STB for the first time. Hence, the client is asked to insert the gender and age of all its family members, including himself. Therefore, every potential STB user in a household has an individual profile.

3.4 User Identification

In TargetAd, user identification may be done explicitly or implicitly. Implicit user identification is achieved without users' input. This kind of identification employs supervised machine learning techniques to try to identify which user is currently watching television. The contents being consumed on a STB serve as the input for a classifier, that, whenever needed, tries to predict which household user is most likely to have seen these contents. This classifier is trained in the server with contents viewed by 10% of television operator's users. The classifier is periodically updated in every STB.

The monitored users, which account to 10% of television operator's user base, must identify themselves explicitly. A user may identify himself explicitly through a graphical interface displayed on the television screen and using STB's remote controller. Moreover, TargetAd is unable to identify several users simultaneously. Therefore, if multiple users want to receive targeted advertisements at the same time, they have to identify themselves explicitly as well.
3.5 Reports

Currently, one of the major drawbacks of television advertising is its lack of accountability. Having TargetAd implicitly the means to implement this feature, it would be inexcusable not to provide this kind of information to advertisers. Therefore, advertisers that choose TargetAd to broadcast their campaigns have the ability to know, for each advertisement, what were the age and gender of its viewers, the percentage of the advertisement that was watched and at what time was the advertisement aired.

3.6 Privacy

Privacy is nowadays a very sound topic. Consequently, preserving TargetAd users’ privacy was something that we bore in mind throughout TargetAd design and implementation. In order to do that, we felt the need to clarify what kind of data is considered personal data. Personal data is defined by European Union directive 95/46/EC as:

‘Personal data’ shall mean any information relating to an identified or identifiable natural person (‘data subject’); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;

The only user data recorded by TargetAd that might be considered as personal data is its users’ age and gender. Since age and gender are two features that, even combined, do not suffice to identify uniquely someone — at least considering a television operator’s user base —, we conclude that TargetAd does not harm its users’ privacy.

4 Implementation

TargetAd consists of several modules implemented independently. All these modules were implemented in Java and run on top of Java Virtual Machine (JVM). The software used during TargetAd’s development is presented below:

1. Java Development Kit (JDK) 6 was used as the development framework;
2. Application Server Java Glassfish;
3. The set of servlets for SIP processing Java SailFin;
4. Eclipse as the Integrated Development Environment (IDE);
5. Ericsson’s Service Development Studio (SDS) for development and testing of IMS applications;
6. MySQL as the database management system;
7. Video LAN Client (VLC) was used as a client and server of video streams;
8. And the suite of software Waikato Environment for Knowledge Analysis (Weka) for implementing, training and testing the classifier.

Additionally, we have also designed and implemented a graphic prototype, in order to simulate the production environment. The prototype is depicted in Figure 2. During TargetAd’s classifier implementation, we had to select the learning algorithm that was best suited to identify implicitly the user who is currently watching television. We’ve tested and assessed several learning algorithms, namely Naive Bayes, Support Vector Machines (SVM), linear regression and decision
trees. Naïve Bayes was the learning algorithm that performed best during the tests, identifying correctly the user 47% of the times. However, we believe that this percentage may be higher once in production, since an incorrect classification may be successfully corrected if the classifier identifies a user whose age and gender do not match any of the household members.

5 Evaluation

After TargetAd’s development, the system was evaluated in order to understand if it achieved the goals initially defined. Hence, three kind of tests were made:

1. Classifier’s success rate was tested;
2. The efficiency of the algorithm used to select and schedule the ads was tested as well;
3. And finally, tests with users were conducted.

As mentioned above, TargetAd’s classifier identifies a user successfully 47% of the times. Nevertheless, it is expected that this value may be higher in a production environment, as previously explained, since an incorrect classification may be corrected dynamically.

In TargetAd, dynamically inserted ads are selected and scheduled through an analogy to the Knapsack Problem. Since this problem is NP-Complete and STBs are devices with limited resources, it became clear that we needed to test this algorithm in order to avail if it was fast enough to select and schedule ads dynamically. For testing purposes, the algorithm was implemented in Python 2.6 and was tested on a computer with an Intel® Pentium™ processor and 128 MB of main memory. Tests showed that for a 10 minute break and handling 100 ads, the algorithm takes almost two seconds. When dealing with shorter breaks, algorithm’s execution time does not exceed one second. In spite of not being a shocking value, television viewers are not used to wait two seconds for a commercial break to start. In fact, we believe that waiting times that exceed one second will be negatively noticed by users. Therefore, future efforts to reduce this algorithm’s execution time must be carried out.

Since improving users’ television experience is one of TargetAd’s goals, it was mandatory to conduct tests with users. In these tests we tried to assess if TargetAd’s targeted advertising presented ads that
were more relevant to users than current television advertising. In order to accomplish this, we displayed
directed and undirected ads to several users, and then asked them to quantify how relevant the displayed
ads were to them. We expected the tested users, in general, to classify the directed advertisements as
more relevant than the undirected ones.
The results of these tests showed that a targeted advertising is 33% more relevant to a user than an
undirected ad, with a confidence of 99%. Consequently, we prove that users' television experience is
better when users are presented with targeted advertising, specifically using TargetAd.

6 Conclusion

In this article we have presented, designed, implemented and evaluated TargetAd, a system for targeting
television advertisements on IPTV. This system has as its main goal presenting relevant ads to its users.

Since some decades ago, we have been witnessing a continuous decrease of television advertising ef-
effectiveness. Nowadays, advertisers are changing their buying habits, advertising more on the internet
and less on television. In fact, online advertising has some advantages over television advertising, as it
provides accountability and enables, at some extent, advertisements targeting.

In order to bring the advantages of internet advertising to television, we have designed, implemented and
evaluated TargetAd, a system for targeting television advertisements on IPTV. This system has as its
main goal presenting relevant ads to its users.

TargetAd is a distributed system, with a module installed on the client-side (STB) and several other
modules installed on the server-side. TargetAd targets advertisements based on users' age and gender.
Moreover, TargetAd allows advertisers to directly submit their advertisements into the system, avoiding
some time-consuming processes currently in use.

In order to TargetAd present targeted advertisements, the system needs to know who is currently watch-
ing television. This implicit user identification is achieved using a Naïve Bayes classifier that, given
television contents as input, it produces a prediction of the sex and gender of the user who has most
likely seen these contents. The selection and scheduling of targeted advertisements is done dynamically
through an analogy to the Knapsack Problem. In this analogy, the commercial break duration is the
knapsack maximum weight, the items are the eligible ads, such as an item's weight is the corresponding
ad's duration and an item's value is given by the number of seconds since the last time the ad was pre-
sented.

TargetAd was submitted to several assessments in which we were able to test several system modules,
namely the bayesian classifier and the algorithm for selecting and scheduling advertisements. Both tests
revealed quite promising results, which justify that TargetAd can eventually become a commercial so-
lution. However, the most important test was the test with users. This test has allowed us to assess if
TargetAd's main goal was actually achieved, as the only way of finding out if the targeted advertisements
presented by TargetAd were more relevant to its users than their undirected counterparts was to ask
the users themselves. We have concluded, from users' feedback, that targeted advertisements were 33%
more relevant to users than undirected ones, justifying that targeted advertising, in this case enabled by
TargetAd, betters users' television experience.

6.1 Future Work

During TargetAd’s development, several ideas and corrections came to our minds as future improvements.
Next we present some of them.
Currently, TargetAd targets advertisements based on users’ age and gender. It would be interesting to include more criteria in the targeting process, in such a way that new enhancements would be driven by advertisers needs. Although the implementation of new targeting criteria do not represent a significant technical challenge, privacy issues must be taken into account as users’ privacy must be always protected.

One of the shortcomings of TargetAd is that the system is unable to implicitly identify several users at the same time. Hence, one of the future improvements of TargetAd — and perhaps the most important one — is the ability to identify multiple users without their interaction. We think that a solution for this problem may be achieved by gathering and analysing more user generated data, like channel navigation patterns, or by using external communications channels, like RFID (Radio-frequency Identification).

Since TargetAd implies a shift in television ads’ spots sales paradigm, we firmly believe that convincing television channels and advertisers of the benefits of TargetAd will be of vital importance in order to TargetAd be accepted as a commercial solution.

References


