Personalized Advertisement-Duration Control for Streaming Delivery

Takashi Oshiba, Yuichi Koike, Masahiro Tabuchi and Tomonari Kamba
NEC Corporation, Internet Systems Research Laboratories
4-1-1, Miyazaki, Miyamae-ku,
Kawasaki, Kanagawa, 216-8555 Japan
+81-44-856-2126

{oshiba@cp, koike@ay, tabuchi@az, kamba@cw}.jp.nec.com

ABSTRACT

This paper describes the development of a streaming advertisement delivery system that controls the insertion of streaming advertisements into streaming content.

Conventional personalization techniques lack a time-control function for advertisement insertion, so the advertisement exposure for each user access can become excessive, much to the annoyance of viewers. This could devalue streaming content by making it less attractive.

In our technique, advertisement insertion control is based on the history of each viewer. This personalization method makes it possible to maintain a balanced ratio of the advertisement length to the content length. As a result, our technique should encourage the growth of Internet streaming services and enable more effective and less intrusive advertising.

Keywords

Internet streaming, Advertisement delivery, Personalization.

1. INTRODUCTION

The use and availability of Internet video streaming services continue to grow with the spread of broadband networks, and streaming advertisement delivery is now attracting considerable attention [1].

A concern with video advertisement delivery, though, is that viewers dislike excessive advertisement insertion into video contents. Persistent and intrusive advertisements are particularly annoying for viewers.

1.1 Web-type Advertisements vs. TV-type Advertisements

The delivery of online advertisements can be classified as Web-type or TV-type delivery.

Web-type online advertisement delivery is typically a conventional banner advertisement inserted [2] into HTML content. In this type of delivery, a banner advertisement and the HTML content are displayed in different areas at the same time. However, viewers who are interested in only the HTML content tend to ignore the banner advertisement, so effective advertisement exposure is difficult to achieve. Using Web-type delivery for streaming advertisements seems unlikely to increase the effectiveness of the advertisement exposure.

On the other hand, in TV-type online advertisement delivery a video advertisement and the video content are sequentially displayed in the same viewing area. We expect TV-type online advertisement delivery to be more effective than Web-type delivery because the advertisement does not have to compete with the desired content for the viewer's attention.

1.2 TV Broadcast Advertisements vs. Streaming Advertisements

In the TV broadcasting industry, the aggregate length of time during which advertisements can be shown is strictly controlled by each TV station. In the U.S.A, the ratio of the length of advertisements to the length of the TV program typically ranges from 25% to 35% [3]. Furthermore, in some countries a limit has been placed on the maximum amount of advertisement exposure. For example, Japanese TV broadcast guidelines [4] limit advertisements to 18% of the length of entertainment content. One reason for limiting advertisement exposure is probably that excessive advertisement exposure would tend to lower the entertainment value of TV programs, thus discouraging potential viewers from watching. With respect to the comparison to Internet advertising, it may also be significant that most TV programs are relatively long, e.g., 30 minutes or more.

For the Internet streaming services industry, the situation is quite different. Although there are some guidelines for rich media advertising (for example, [5]), these guidelines do not currently provide any policy related to the length of streaming advertisements.

Furthermore, managing the aggregate length of delivered advertisements is difficult because:

- Each viewer receives different content from many sources and the personal access pattern will depend on the viewer's preferences and environment.
- The number of Internet content sources greatly exceeds the number of TV channels.

- Internet content varies greatly in length, for example, ranging from 30 seconds to several minutes. Making short, semi-short, semi-long, and long versions of the same advertisement to fit the varying lengths will be expensive and troublesome for sponsors.

1.3 Excessive Advertisement Exposure

Because of the lack of industry guidelines and these difficulties regarding the content length, there is a danger of excessive advertising devaluing the attractiveness of streaming content.

For example, imagine a streaming media portal website that offers many short news clips, music clips, and sports highlights, each lasting a few minutes or less than a minute. If a streaming advertisement delivery system inserted an advertisement longer than the entire short video clip, viewers would probably find this objectionable.

Thus, we think a TV-type streaming advertisement delivery system should have a function that prevents excessive advertisement exposure [6][7]. Specifically, we think the time length of inserted advertisements should be controlled for each viewer so that a viewer who sequentially watches content from many streams will not be overwhelmed by the delivery of advertisements. In other words, the advertisement insertion should be personalized for each viewer.

In this paper, we propose such a personalization technique for streaming advertisement delivery. This technique governs advertisement exposure based on each viewer's usage data. The amount of advertising inserted will therefore be proportional to the amount of content delivered. We have implemented a prototype streaming advertisement delivery system that uses this personalization technique.

2. PERSONALIZATION TECHNIQUES

2.1 Classification of Conventional Approaches

Personalizing the Web experience for a user is the holy grail of many Web-based applications. Several mathematical techniques are used in conventional personalization approaches [8][9][10].

- **Nearest-neighbor collaborative filtering algorithms** [11][12][13] compute the distance between consumers based on their preference history. These algorithms can be used, for example, to predict how much a customer will like a product. Predictions are computed by taking the weighted average of the opinions of a set of nearest neighbors for that product. By using these predictions, a product likely to appeal to a customer can be recommended. In e-commerce and one-to-one marketing, recommended products for each customer can be used for cross-selling and up-selling. Recommendation is one personalization technique.

- **Data mining** is used to extract the characteristic tastes of customers, and to analyze a customer's potential tendencies. WebWatcher [14], SiteHelper [15], and Letizia [16] learn the preferences of customers. WebWatcher keeps track of a customer's web surfing, and identifies potentially interesting links for the customer. SiteHelper analyzes a customer's web access history, and recommends other web pages that the customer has not visited. Letizia analyzes a customer's web access history and the customer's bookmarks, and displays similar web pages for each customer.

- **Linear programming** is used in ADWIZ [17][18] to maximize the click-through rate for a banner advertisement. Depending on the search keyword supplied by a customer to a search engine, ADWIZ delivers the best matching banner advertisement, thus displaying advertisements that match a customer's immediate interests.

2.2 Shortcoming of the Conventional Approaches

These conventional approaches rely on the consumer's usage information to personalize in that they analyze customer purchase histories, web server access logs, customer bookmarks, search keywords, etc. However, none adequately takes into account the length of time a user spends viewing any particular content. <u>**Therefore, they cannot be used to manage the appropriate aggregate length of advertisement insertions.**</u> In a manner of speaking, they are best categorized as forms of frequency-based or genre-based personalization.

The problem that arises is that a short video clip might always contain a relatively long advertisement that a viewer must sit through if the conventional approaches for personalization are applied to a streaming advertisement delivery system because these approaches are not sensitive to the advertisement or content length. <u>**They lack a duration-control function for advertisement insertion.**</u> The conventional approaches for personalization will therefore not help us prevent excessive advertisement exposure.

3. PERSONALIZED ADVERTISEMENT-DURATION CONTROL

We propose the technique of personalized advertisement-duration control – a personalization technique based on the passage of time.

Our technique records the accumulated time length of both the streaming content and the streaming advertisements delivered to a user. It also records the accumulated number of times any particular streaming content or advertisement is delivered. The accumulation data regarding the length of time and the number of times are stored separately as the **usage data** of each viewer.

Our technique also uses several preset **control data**; i.e., values used to control the personalization processes. The length of the delivered content and that of the delivered advertisements can be balanced by using both the usage and the control data.

Our personalization technique consists of three distinctive processes: **ad-insertion decision**, **ad-length decision**, and **real-time updating of usage data** (Figure 1). Through these processes, <u>**our technique provides a duration-control function for ad insertion.**</u>

- **The ad-insertion decision** process dynamically decides whether a streaming advertisement should be inserted into streaming content.

- **The ad-length decision** process selects a streaming advertisement based on the length of the streaming content requested by a viewer. The ratio of the content length to the selected advertisement's length can be controlled by using both the usage data and the control data.

- **The real-time updating of usage data** process keeps the viewer usage data up-to-date through real-time feedback.

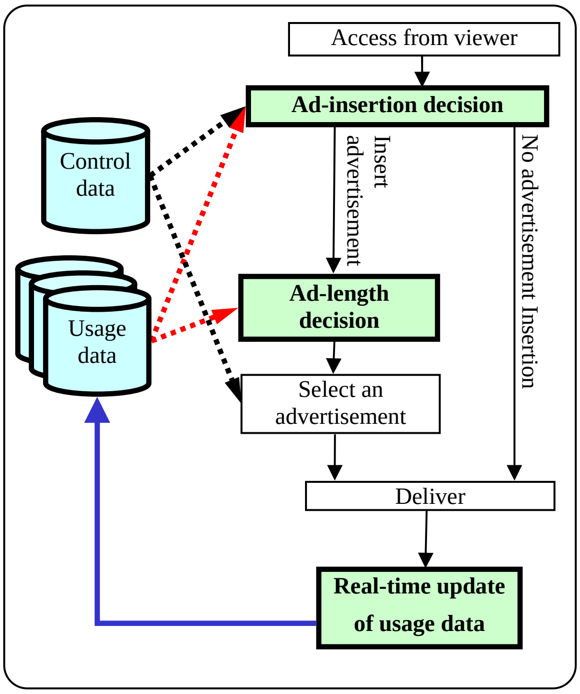

Figure 1: Flowchart of our technique.

3.1 Usage Data and Control Data

We will explain the detailed mechanism of our technique in terms of each parameter of both the usage data and the control data. We assume that the range of each parameter value is greater than zero.

In our technique, the usage data consists of the parameters:

- *progSec*: total length of streaming content that a viewer has seen.
- *progTimes*: total number of times the viewer has seen particular streaming content.
- *adSec*: total length of streaming advertisements that the viewer has seen.
- *adTimes*: total number of times the viewer has seen a particular streaming advertisement.

These parameters are dynamically updated.

The control data consists of the parameters:

- *secRatioGoal*: target value of *adSec / progSec*.
- *timesRatioGoal*: target value of *adTimes / progTimes*.
- *secRange*, *timesRange*: threshold values used to revise the difference between target values and actual results.
- *adBlur*: a value used for the above revision.

These control parameters are preset and are essentially static data.

3.2 Ad-Insertion Decision Process

This process is invoked when a viewer requests access to streaming content. The detailed process is as follows:

```
If ( timesRatioGoal – ( adTimes / progTimes )
        < –timesRange )
   // (1) actual result is too large compared to target value.
   return NO_WITHOUT_AD;
If ( adSec / progSec < secRatioGoal )
   // (2) an advertisement must be inserted into the content.
   return YES_WITH_AD;
else
   // (3) no advertisement is inserted.
   return NO_WITHOUT_AD;
```

An advertisement will be inserted into the content in case (2), but no advertisement is inserted in case (3) (Figure 2).

In case (1), the ratio of the number of times content is delivered to the number of times an advertisement is delivered significantly exceeds the target value. In this case, regardless of cases (2) or (3), no advertisement is inserted.

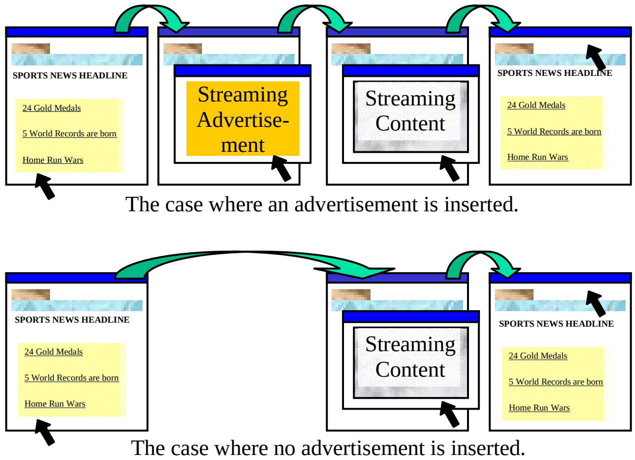

Figure 2: Ad-insertion decision process.

3.3 Ad-Length Decision Process

If an advertisement is to be inserted, the ad-length decision process is invoked.

In our technique, advertisements are grouped according to their length with each group having a representative time length.

With *progLength* as the content time length, the detailed process is as follows:

```
gap = secRatioGoal – adSec / progSec;
If ( gap < –secRange )
   // (4) actual result is too large compared to target value.
   return progLength * secRatioGoal – adBlur;
else if ( –secRange < gap && gap < secRange )
   // (5) actual result is near  target value.
   return progLength * secRatioGoal;
else if ( secRange < gap )
   // (6) actual result is too small compared to target value.
   return progLength * secRatioGoal + adBlur;
```

By using the return-value, the advertisement group from which an advertisement can be inserted into the content is selected. That is, the advertisement group whose representative time length is nearest to the return-value is selected. Each advertisement belonging to the selected advertisement group becomes a candidate for insertion. The actual advertisement to be inserted is chosen at random from among the candidates.

In cases (4) and (6), the ratio of the delivered content length to the delivered advertisement's length is significantly different from the target value.

Grouping advertisements by their time length is a way to ensure fairness in the advertisement selection even if there are several advertisements of almost the same length.

3.4 Real-time Updating of Usage Data

The change in each parameter in response to an access request from the viewer is reflected in the usage data. The ad-insertion processes remain accurate even if a viewer sequentially views many content streams because the updating process is executed whenever the viewer accesses any content.

Through these three processes, personalized streaming advertisements can be delivered to each viewer. The personalization processes are controlled by *secRatioGoal* and *timesRatioGoal*.

4. IMPLEMENTED SYSTEM

We have implemented a streaming advertisement delivery system that uses our personalization technique. Figure 3 shows a prototype of a streaming media portal site connected to our streaming advertisement delivery system.

When a viewer clicks a news title on the web site, our personalization processes are executed and create a play list file which contains the URL of a streaming advertisement and the URL of streaming content. The play list file is transmitted to the viewer. In the viewer's computer, a media player automatically starts to play back the streaming advertisement and the streaming content. (Our system currently supports Windows Media Player and RealPlayer.)

Figure 3: Snapshot of the streaming media portal web site.

Our streaming advertisement delivery system manages the control data and the usage data shown in Figure 1 by using a database management system (DBMS), i.e., Oracle 9i [19], Microsoft SQL Server [20], PostgreSQL [21], MySQL [22], etc. An administrator of our system can change the parameters of the control data by using a web-based user interface.

Each viewer is identified by our system through an HTTP cookie. Our system also has an option to identify viewers by obtaining the IP address of each viewer's computer. After the identification, our system assigns a viewer ID to each viewer. The DBMS stores the viewer ID into the usage-data database. With this mechanism, our system can recognize the relationship between a particular viewer and the particular record of usage data.

4.1 Pluggable System Architecture

Our streaming advertisement delivery system has a plug-in API (application programming interface) so that it can be connected to and cooperate with other advertisement delivery engines. (Of course, our system can also work without the help of other engines.)

The plug-in API has arguments concerning:

- Streaming content metadata
- Viewer metadata
- The length of the advertisement selected in the ad-length decision process.

The streaming content metadata is, for example, the content genre, the content length, and the content author. The viewer metadata is, for example, the viewer IP address and the web browser information. If no advertisement is inserted, the advertisement length is set to zero.

If our system connects to an advertisement delivery engine that uses genre-based personalization, our system can add our personalization function to that engine (Figure 4). Hybrid personalization could thus be realized through cooperation between our system and another advertisement delivery engine.

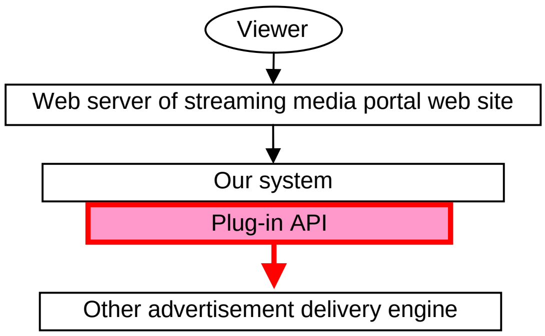

Figure 4: Plug-in API.

4.2 System Architecture for Streaming Advertisement Delivery

Figure 5 shows the system architecture for streaming advertisement delivery. Because a play list file created by our system contains the URLs of a streaming advertisement and a streaming content item, our system does not need to hold its own streaming media server; i.e., our system and the streaming media server can be deployed in separate network segments. Streaming media servers for streaming advertisement and those for streaming content can also be deployed in separate network segments. Thus, streaming advertisement delivery across multiple web sites can be done by using our system.

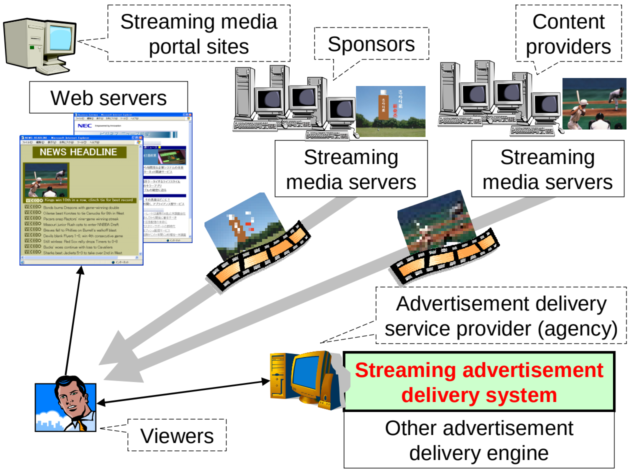

Figure 5: System architecture for streaming advertisement delivery business.

5. CHANGE IN THE USAGE DATA

5.1 Initial Examination

We prepared 12 video clips of sports news as streaming content. Each clip was about 35 seconds long. We also prepared 13 video clips of streaming advertisements for merchandise. There were two advertisement groups – four advertisements were between 1 and 2 seconds long, and nine advertisements were between 5 and 7 seconds long.

A person in our research group became a viewer as the test subject in this examination. The viewer randomly watched various content items (and advertisements) sequentially through a streaming media portal site. The viewer selected video titles nine times.

Figure 6 shows how the viewer's usage data changed. (The horizontal broken line shows the value of *secRatioGoal*.)

In the ad-insertion decision process, an advertisement was inserted into the content if the value of *adSec / progSec* was below the *secRatioGoal* line. Otherwise, no advertisement was inserted.

The lightly shaded area above and below the line represents twice the value of *secRange*. If the value of *adSec / progSec* was inside this area, case (5) of the ad-length decision process (Section 3.3) was executed. Otherwise, case (4) was executed if the value of *adSec / progSec* was above the area, and case (6) was executed if the value was below the area.

We preset the control data as follows:

- *secRatioGoal* (target value of *adSec / progSec*) = 7.5%
- *timesRatioGoal* (target value of *adTimes / progTimes*) = 60.0%
- *secRange* (threshold value) = 1.0 s
- *timesRange* (threshold value) = 20.0 points.
- *adBlur* (used for revision) = 3.0 s.

Figure 6 shows that the value of *adSec / progSec* converged on the value of *secRatioGoal*, thus demonstrating that the behavior of the streaming advertisement insertion was effectively managed by the control data with our technique.

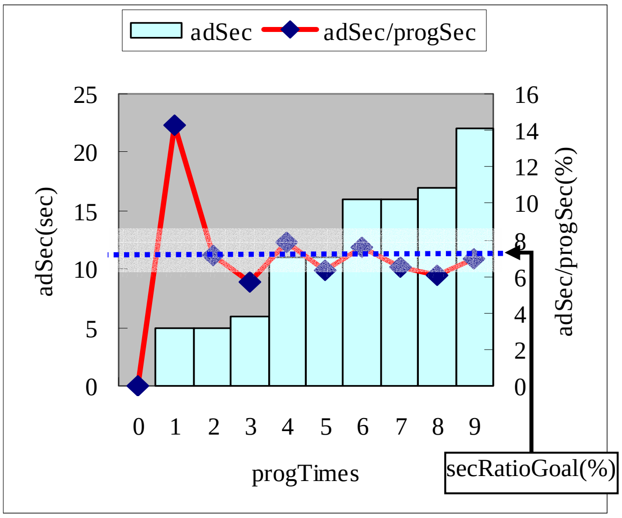

Figure 6: Change in usage data.

With reference to the Figure 6 results, we then varied the control parameters and the lengths of the advertisements and content to examine how this affected the application of our technique. The same person was again the test subject and randomly selected nine content items. Each examination of variation was run just once.

5.2 Variation of the Time Length Target Ratio

First, we varied the time length target ratio (*secRatioGoal*) while keeping the other control parameters the same as in Section 5.1.

This control parameter should be of particular interest to both sponsors and the service providers who deliver advertisements because it directly influences the total time length of delivered advertisements.

We started by increasing the value of *secRatioGoal:*

- *secRatioGoal* = 10.0%.

Figure 7 shows that the insertion of advertisements from the longer advertisement group (i.e., between 5 and 7 seconds) increased. Again, the value of *adSec / progSec* converged with the value of *secRatioGoal*.

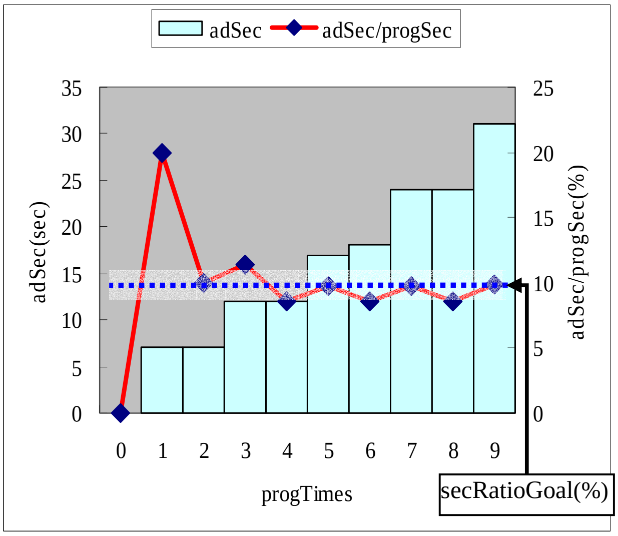

Figure 7: secRatioGoal increased to 10.0%.

We then lowered the value of *secRatioGoal:*

- *secRatioGoal* = 3.0%.

In this case, the insertion of advertisements from the shorter advertisement group (i.e., between 1 and 2 seconds) increased (Figure 8), and the value of *adSec / progSec* still converged with the value of *secRatioGoal*.

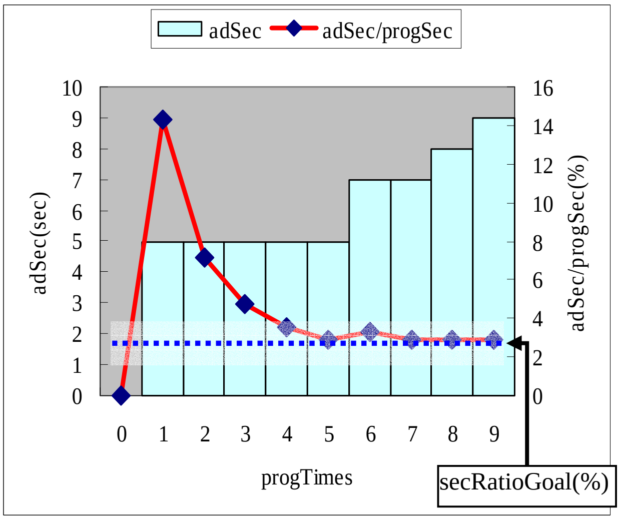

Figure 8: secRatioGoal decreased to 3.0%.

5.3 Variation of the Advertisement Length

Next, we increased the variation in the advertisement length. We prepared various advertisements of differing lengths; e.g., 1, 3, 5, 7, or 9 seconds. We created five advertisement groups and assigned each of 25 advertisements to a group. The control parameters and the total number of content items were the same as in Section 5.1.

Sponsors are likely to be interested in the result of this change because preparing advertisement versions of various lengths will directly affect the production cost of advertisements.

Figure 9 shows convergence at *secRatioGoal* after the content was viewed only two times, which was much faster than the five times needed for convergence in Figure 6,Figure 7, and Figure 8.

This result shows that preparing advertisement versions of various lengths could enable a more detailed level of personalization, although at a higher advertisement production cost.

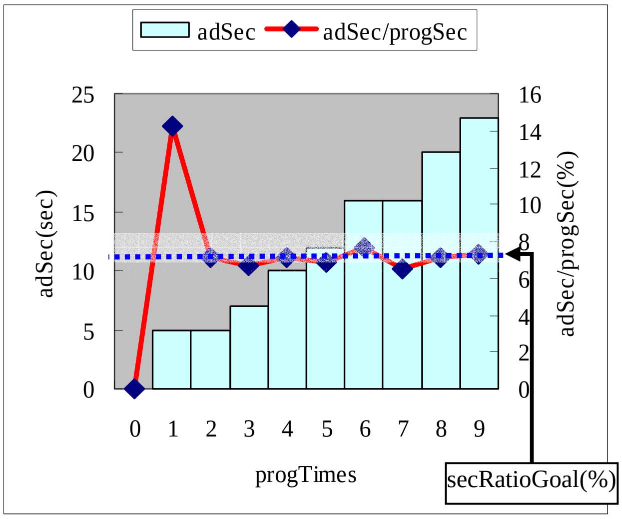

Figure 9: Variation in the advertisement time length.

5.4 Variation of the Content Length

We then increased the variation in the time length of content items. We prepared 30 content items of various lengths; e.g., 20, 35, 40, 50, 60, or 100 seconds. The control parameters and the total number of advertisements were the same as in Section 5.1.

Content owners should be interested in the result of this change because the length of advertisements inserted into the content could affect the attractiveness and ultimate value of the content.

As before, the value of *adSec / progSec* converged with the value of *secRatioGoal* (Figure 10).

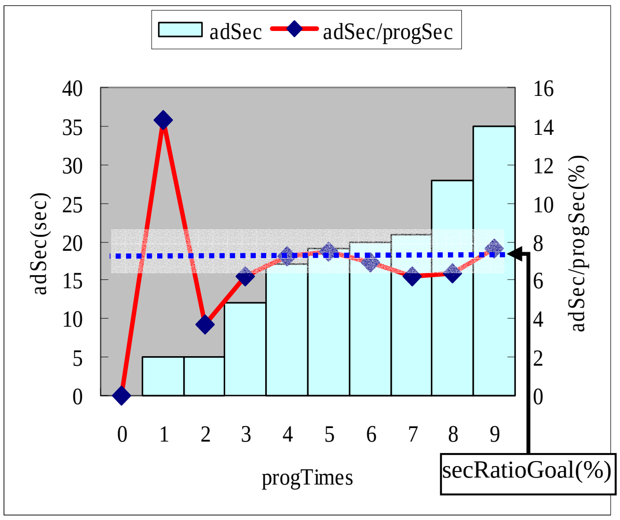

Figure 10: Variation in the content time length.

5.5 Discussion

From a long-term point of view, excessive advertisement exposure was prevented in all cases examined above. We thus confirmed that our technique provides an effective time-control function that can be used to personalize streaming advertisement insertion.

From a short-term point of view, varying the length of advertisements affects the convergence speed. There is thus a trade-off between this variation and the detail level of personalization. However, when a viewer sequentially accesses many content items, the result will be almost the same.

On the other hand, varying the content length has little effect on the convergence speed.

6. CONCLUSION

In this paper, we have explained how our personalization technique can be applied in a streaming advertisement delivery system to prevent the devaluation of streaming content due to excessive advertisement exposure.

Our approach should be effective for video advertisement delivery to mobile phones or other mobile terminals, as well as for broadband Internet streaming. There is a strict limit on the file-size of video clips that can be played back on current mobile terminals, and excessive exposure to advertisements can easily occur with such short video clips.

6.1 Future Work

We will continue to develop our system in a way that will allow it to connect to ADWIZ [17][18] to realize a form of hybrid personalization. We also plan to extend the advertisement selection function to enable insertion of multiple advertisements. We will need to investigate ways to schedule multiple advertisements and to avoid advertising competing products in the same timeslot [23].

We will create a model of user viewing processes [24] by studying the viewing habits of test subjects. In addition, an evaluation of each control parameter is also needed.

7. REFERENCES

[1] Yankee Group News Release, June 18, 2001, “Yankee Group Estimates Spending on Streaming Media–Enabled Advertising Online to Grow to $3.1 Billion by 2005.”, http://www.yankeegroup.com/

[2] DoubleClick. http://www.doubleclick.com.

[3] American Association of Advertising Agencies (AAAA) and Association of National Advertisers, Inc. (ANA). 2001 Television Commercial Monitoring Report, Online version is available at http://www.ana.net/com/ta/clutter.pdf

[4] National Association of Commercial Broadcasters in Japan. Broadcasting Standards (part XVIII. Standards of Broadcasting Time of Advertising). Online English version is available at http://www.nab.or.jp/htm/english/nabstd99.PDF.

[5] Interactive Advertising Bureau (IAB). Voluntary Guidelines for Rich Media Ad Formats. Online version is available at http://www.iab.net/news/content/rich_media.html.

[6] T. Oshiba, K. Nakajima, Y. Koike, M. Tabuchi, and T. Kamba. Personalized Streaming Advertisement Delivery through Internet Broadcasting Services. In *64th Annual Conference of Information Processing Society of Japan*, vol.3, pp.59-60, 2002 (in Japanese).

[7] T. Oshiba, Y. Koike, M. Tabuchi, and T. Kamba. Temporal-Axis Personalization Techniques for Streaming Advertisement Delivery on Internet Broadcasting Service. In *2nd Study Group Notes of Broadcast Communication & Computing (BCC) of Information Processing Society of Japan*, No.BCCgr-2, pp.25-31, 2002 (in Japanese).

[8] J. Srivastava, R. Cooley, M. Deshpande, and P. N. Tan. Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data. In *SIGKDD Explorations*, 1(2): pp.12-23, 2000.

[9] J. B. Schafer, J. A. Konstan, and J. Riedl. E-Commerce Recommendation Applications. In *Data Mining and Knowledge Discovery*, vol.5, pp.115-153, 2001.

[10] A. Kobsa, J. Koenemann, and W. Pohl, Personalized hypermedia presentation techniques for improving online customer relationships. *The Knowledge Engineering Review*, Vol.16:2, pp.111-155, 2001, Cambridge University Press.

[11] Broadvision. http://www.broadvision.com.

[12] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl. Grouplens. An open architecture for collaborative filtering of netnews. In *Proceedings of ACM CSCW'94 Conference on Computer-Supported Cooperative Work*, pp. 175-186, 1994.

[13] U. Shardanand and P. Maes. Social information filtering: Algorithms for automating “word of mouth”. In *Proceedings of ACM CHI'95 Conference on Human Factors in Computing Systems*, pp.210-217, 1995.

[14] T. Joachims, D. Freitag, and T. Mitchell. WebWatcher: A tour guide for the World Wide Web. In *15th International Conference on Artificial Intelligence*, 1997.

[15] D. S. W. Ngu and X. Wu. SiteHelper: A localized agent that helps incremental exploration of the World Wide Web. In *6th International World Wide Web Conference*, 1997.

[16] H. Lieberman. Letizia: An agent that assists web browsing. In *Proc. of the 1995 International Joint Conference on Artificial Intelligence*, 1995.

[17] M. Langheinrich, A. Nakamura, N. Abe, T. Kamba, and Y. Koseki. Unintrusive Customization Techniques for Web Advertising. In *8th International World Wide Web Conference*, 1999.

[18] A. Nakamura. Improvements in Practical Aspects of Optimally Scheduling Web Advertising. In *11th International World Wide Web Conference*, 2002 (to appear).

[19] Oracle 9i Database. http://www.oracle.com/ip/deploy/database/oracle9i/

[20] Microsoft SQL Server 2000. http://www.microsoft.com/sql/default.asp

[21] PostgreSQL. http://www.postgresql.org/

[22] MySQL. http://www.mysql.com/

[23] K. Hagele, C. Dunlaing, S. Riis. The Complexity of Scheduling TV Commercials. In *Proceedings of First Irish*

Conference on the Mathematical Foundations of Computer Science and Information Technology, 2000.

[24] T. Syeda-Mahmood, D. Ponceleon. Learning Video Browsing Behavior and Its Application in the Generation of Video Previews. In *Proceedings of 9th ACM International Conference on Multimedia*, pp.119-128, 2001.