Individualizing User Profile from Viewing Logs of Several People for TV Program Recommendation

Koichi Iguchi
Graduate School of Engineering Science
Osaka University
Osaka, Japan
iguchi@nishilab.sys.es.osaka-u.ac.jp

Yoshinori Hijikata
Graduate School of Engineering Science
Osaka University
Osaka, Japan
hijikata@sys.es.osaka-u.ac.jp

Shogo Nishida
Graduate School of Engineering Science
Osaka University
Osaka, Japan
nishida@sys.es.osaka-u.ac.jp

ABSTRACT

Recommendation services for e-commerce and content distribution sites are increasing in popularity. An individual is identified by a login certification, and then information pertaining to their interests and preferences are collected. Finally, the recommendation service recommends items to the user based on the information gathered. Recommendation services are not as popular for recommending television programs as they are for Web-based services. The primary reason for this discrepancy is that it is difficult to identify an individual sitting in front of a television, which is usually shared by several members of a family. Viewers generally do not want to perform additional operations on the remote control for login authentication rendering it impossible to create a user profile that contains personal interests and preferences. Therefore, in this study, we propose a technique to individualize a user profile from the combined viewing logs of several television viewers.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Information filtering; H.3.4 [Systems and Software]: User profile and alert services

General Terms

Algorithms, Experimentation, Human Factors

Keywords

recommender systems, personalization, user profile, TV program, content-based filtering, time interval

1. INTRODUCTION

Television (TV) programs have diversified since the beginning of digital terrestrial broadcast. As a result, more choices are available to viewers. However, despite having a wide variety of programs to choose from, it is difficult for viewers to find programs they actually want to watch. Research on TV program recommendations has been conducted to solve this problem. Many of the proposed systems, however, recommend programs to the whole family rather than to individual family members. Very few systems exist that recommend TV programs to a specific user when they share TV set with multiple people.

Recommendations for individuals are realized in many e-commerce and content distribution sites. In these services, the system identifies an individual by a login certification. Users are generally required to enter their user name, identification (ID), and password. Login certification is not popular for TV viewers because many viewers do not want to perform additional login operations on a remote control. A TV set is usually shared by several people. When the system recommends TV programs based on viewing history without a login certification, it provides a recommendation list fit for the whole family not an individual. This can be frustrating when one member of the family chooses to watch TV by themselves.

In this study, we propose a technique that recommends TV programs to an individual. This technique assumes each family maintains a habitual viewing pattern with regards to both the time they turn on the TV and the content (TV program) they view. We also assume there exists a dominant user who selects a TV programs during each viewing hour. For example, if one family member watches cartoons every day at 10 am and another watches the news every day 12 pm. This enables the system to individualize each user based on the collected hourly viewing patterns. We also consider the day of the week in our proposed technique. One of the primary advantages of the proposed technique is that it does not require login certification.

Many viewers have a specific program they watch habitually (hereafter referred to as a habitual program). For example, a user watches a particular drama every Sunday at 8 am. In this case, if the drama is recommended to them, the recommendation is of no value because they already plan to watch it. However, the user might turn on the TV even when there is no habitual program on air just because they have nothing to do during the hour. In this case, they rapidly switch channels just after the TV is turned on. Another case is when the habitual program is not airing during its regularly scheduled time. This usually occurs during a season of

IMCOM'15, January 08-10 2015, BALI, Indonesia
Copyright 2015 ACM 978-1-4503-3377-1/15/01 ...$15.00.
http://dx.doi.org/10.1145/2701126.2701132.

TV program reorganization when special TV programs are on air. Particular emphasis is placed on the above situations. When recommendations are made for a specific time, they are strongly influenced by the programs in the viewing logs; this is especially true if viewing logs from only a specific time interval on a specific day of the week are used. This makes it difficult to generalize a user's preferences and create a personal profile. Therefore, we combine days of the week with similar viewing patterns into one group to create a user profile.

The Electronic Program Guide (EPG), which contains descriptions of TV programs, is used to recommend TV programs. A content-based filtering technique [7] is used as a recommendation mechanism. Users receive recommendations based on individual interests without any explicit input operations.

The remainder of this paper is organized as follows. Section 2 discusses related works, and the proposed method is introduced in Section 3. Section 4 describes the data set used in this study, and Section 5 presents the results of the proposed method's evaluation. Finally, concluding remarks are provided in Section 6.

2. RELATED WORK

Studies of user profiling for recommender systems in the TV domain are based on either user preference extraction or user profile individualization. User profile individualization can be classified into two types: individualization for an individual and individualization for a group.

First, we review literature related to user preference extraction. Ali et al. proposed a collaborative filtering (CF) system called TiVo, which is implemented for a video recorder [1]. Moreover, they compensated for the sparseness [14] of a collaborative filtering rating matrix by using evaluation values submitted by users. Bellekens et al. improved the quality of user profiles by using an individual's search queries or evaluation values for TV programs submitted through mobile devices and semantic information on the Web [3]. Nakamura et al. proposed a system that extracts user preferences using life logs (Web browsing history and global positional system (GPS) information) in addition to viewing histories [12]. Takama et al. proposed a system that recommends TV programs using emotions estimated from reviewers' comments of TV programs [15]. All of these studies address the acquisition of user preferences. Furthermore, these studies require users to perform additional input operations through various devices.

Second, we review studies related to user profile individualization. Zhang et al. created an individual user profile by splitting evaluation values given by several users [17]. They vectorized these values and applied them for using a subspace clustering technique [10] to estimate the number of users and their preferences. However, each user must submit an evaluation value for each item to receive recommendations. Systems that recommend TV programs for a group of viewers also exist. These systems assume that a TV is shared and watched by several people. Masthoff investigated the relationship between the preferences of three users belonging to the same group [11]. Masthoff showed that the highest level of satisfaction is achieved when the system reflects user preferences equally in the user profile. Seko et al. proposed a method to balance the preferences of two users based on their individual evaluation values and viewing frequencies [13]. Zhiwen et al. considered how to optimize TV programs included in the recommendation list [16]. They proposed merging several user profiles identified through login certification. All of these studies make recommendations for a group of users rather than an individual.

3. APPROACH AND PROPOSED METHOD

3.1 Approach

We assume most families follow a habitual pattern of TV watching. For example, on weekdays, a father watches the news early in the morning before his commute, a mother watches a drama during the daytime hours, and their children watch anime films in the evening. After work, the father comes home and watches the news again. On weekends, all members of the family watch a comedy show in the morning, and then the father watches a sports program in the daytime. No one watch TV in the late afternoon, and the family watches a movie at night. Each member of the family maintains a habitual pattern of viewing behavior that is by the family's lifestyle. Thus, we assume a pattern exists in the TV viewing hours and the content of the TV programs selected in a given hour. We individualize a user profile using this habitual pattern and recommend TV programs for an individual.

In particular, we estimate the time interval when the member(s) of the family are likely to watch TV. Since we assume a dominant user controls the TV in each time interval, we recommend TV programs to this user. However, the dominant user may vary between weekdays and weekends. This problem can be solved if the individual days of the week are considered rather than the week as a whole. Recommendation results are strongly influenced by TV programs that are watched habitually (habitual programs). Therefore, we cluster days with similar viewing patterns and estimate a time interval for each cluster. Features of the TV programs watched in this estimated interval are extracted using EPG; thus, enabling a user profile to be created that reflects the general interests or preferences of the user (independent from usual TV programs). When a user accesses this system, they receive a recommendation for TV programs currently broadcasted using the user profile of the time interval.

3.2 User Profile Individualization

This subsection describes the proposed method for creating an individualized user profile in greater detail. A time interval when the TV is likely to be turned on is defined as an active interval. The viewing log records the start time, end time, and name of the TV program watched. The proposed method groups together days with similar viewing patterns and estimates the active interval for each pattern.

3.2.1 Estimating Days with Similar Viewing Behaviors

Active intervals may differ for each family according to the day of the week; this is largely influenced by the cycle of working and non-working days. In general, working days are Monday through Friday and non-working days are Saturday and Sunday. The proposed method automatically estimates days with similar viewing behaviors since some of the family's non-working days are not Saturday and Sunday. The details of this process are described below.

Viewing logs over the course of W weeks are used to create a user profile. Figure 1 explains how to create a representative ON/OFF log for day d, where $d \in DAY = \{Monday, Tuesday, ..., Sunday\}$, using logs from the same day of the week over the course of W weeks. A log contains information about when the TV is turned on or off on a given day; specifically, we denote an ON/OFF log by o, where t_{on} is the time when the TV is turned on and t_{off} is the time when the TV is turned off. The ith ON/OFF log is denoted by $o_{d,w}(i)$, where $I_{d,w}$ is the total number of ON/OFF logs on day d in the wth week ($1 \leq w \leq W$). The complete set of $I_{d,w}$ ON/OFF logs on day d in the wth week is denoted by $O_{d,w}(i)$. We also generalize a start and end time for the set of ON/OFF logs on day d. Generalized ON/OFF logs are called representative ON/OFF logs, and the set of representative ON/OFF logs on day d is denoted by O_d^{Rep}. The jth representative ON/OFF log is denoted by $o_d^{Rep}(j)$, where J_d is the total number of representative ON/OFF logs on day d.

Algorithm 1 creates a set of representative ON/OFF logs. ON/OFF logs that include a common time interval are extracted from the set of ON/OFF logs $O_{d,w}$ for W weeks. An extracted log is denoted by $S_d(j)$, and entire set of $S_d(j)$ logs is denoted by S_d. The algorithm scans 24 hours a day in m minute intervals. The scanning time is denoted by t, where $t \in TIME$, a set of scanning time. ON/OFF logs $o_{d,w}(i)$ on day d of the wth week that include t are extracted from $O_{d,w}$; the extracted logs are denoted by $S_d(j)$ (lines 5 to 7). The same operations are repeated for every scanning time to obtain the set S_d. Representative times $t_{on}^{Rep}(j)$ and $t_{off}^{Rep}(j)$ (generalizations of t_{on} and t_{off}) when the TV is turned on and off, respectively, are calculated using one of three generalizing methods. The representative ON/OFF log $o_d^{Rep}(j)$ that includes $t_{on}^{Rep}(j)$ and $t_{off}^{Rep}(j)$ is obtained at line 13. This procedure is repeated to obtain the set O_d^{Rep} of representative ON/OFF logs (line 17).

Outlier ON/OFF logs may exist. An outlier ON/OFF log is a log that is recorded during a time when a user does not ordinarily watch TV or when they watch an exceptionally long TV program. We use one of the following three generalization methods to obtain $t_{on}^{Rep}(j)$ and $t_{off}^{Rep}(j)$: the average ($Average$), median ($Median$), and selective average ($Average^S$). The average method calculates representative values; however, it is vulnerable to outliers. The median method utilizes basic statistical methods, and unlike the average method, is insusceptible to outliers. Selective averaging obtains an average after removing outliers using the Smirnov-Grubbs test [8]. This test is a general statistical method for detecting outliers.

The set O_d^{Rep} of acquired representative ON/OFF logs is vectorized. The top portion of Fig. 2 shows representative ON/OFF logs for each day. A representative ON/OFF log vector for day d is denoted by $\boldsymbol{v_d}$, and its element during scanning time t' on day d is denoted by $v_d(t')$.

Algorithm 2 vectorizes representative ON/OFF logs. This algorithm scans for 24 hours a day at m' minute intervals; the scanning time is denoted by t', where $t' \in TIME'$ (the set of scanning times $TIME'$ can be identical to $TIME$). In a representative ON/OFF log $o_d^{Rep}(j)$ on day d, if $t_{on}^{Rep}(j) \leq t' \leq t_{off}^{Rep}(j)$, then $v_d(t') = 1$; otherwise, $v_d(t') = 0$ (lines 5 to 9). These operations are repeated for $TIME'$, and a representative ON/OFF log vector $\boldsymbol{v_d}$ is obtained (bottom half of Fig. 2). After the acquisition of $\boldsymbol{v_d}$ for all days, the representative ON/OFF log vectors are segmented into X clusters. K-means [4] is used as the clustering method. individual clusters $DAY^c(x)$, where $1 \leq x \leq X$, are generated for days that have similar ON/OFF logs. The complete set of clusters is denoted by DAY; note that $DAY^c(1) \cup DAY^c(2) \cup ... \cup DAY^c(X) = DAY$.

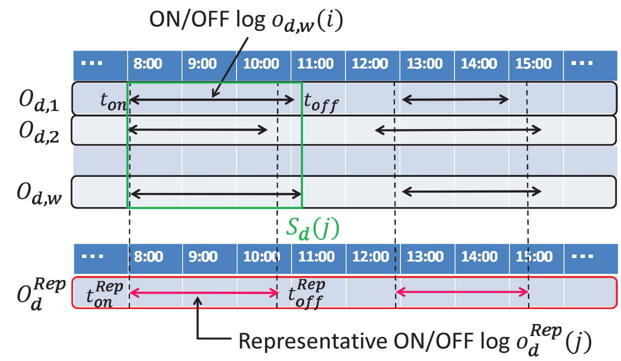
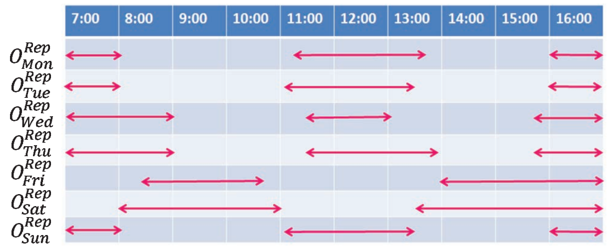


Figure 1: Generating times when the TV is turned on and off.



	7:00	8:00	9:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00
$\boldsymbol{v_{Mon}}$	1	0	0	0	1	1	1	0	0	1
$\boldsymbol{v_{Tue}}$	1	0	0	0	1	1	1	0	0	1
$\boldsymbol{v_{Wed}}$	1	1	0	0	1	1	0	0	0	1
$\boldsymbol{v_{Thu}}$	1	1	0	0	1	1	1	0	0	1
$\boldsymbol{v_{Fri}}$	0	1	1	1	0	0	0	1	1	1
$\boldsymbol{v_{Sat}}$	0	1	1	1	0	0	1	1	1	1
$\boldsymbol{v_{Sun}}$	1	0	0	0	1	1	1	0	0	1

Figure 2: Vectorization of representative ON/OFF logs.

3.2.2 Estimating Active Intervals

After generating clusters of days using the method explained in the previous subsubsection, active intervals are estimated for each cluster. Figure 3 shows generated active intervals for representative ON/OFF logs in each cluster. For a cluster x, $1 \leq x \leq X$, a set ACT_x of active intervals $act_x(j)$ is obtained using the same operation as Algorithm 1 by replacing $O_{d,w}$ with O_d^{Rep} (recall $I_{d,w}$ is the number of representative ON/OFF logs for day d in week w). Note that $act_x(j)$ and ACT_x correspond to $o_d^{Rep}(j)$ and O_d^{Rep} in Algorithm 1 respectively. The operations are repeated for every cluster. A set of active intervals is obtained for each cluster.

Algorithm 1 Creation of representative ON/OFF logs.

for all $d \in DAY$ do
  initialize an internal variable j;
  for all $t \in TIME$ do
    for $w = 1$ to W do
      for $i = 1$ to $I_{d,w}$ do
        append $o_{d,w}(i)$ to $S_d(j)$ when it includes scanning time t;
      end for
    end for
    $j \Leftarrow j + 1$;
  end for
  get the number of elements J_d in S_d;
  for $j = 1$ to J_d do
    create $o_d^{Rep}(j)$ by $Average$, $Median$, $Average^S$;
    append $o_d^{Rep}(j)$ to O_d^{Rep};
  end for
end for
return O_d^{Rep};

Algorithm 2 Vectorization of representative ON/OFF logs.

$\boldsymbol{v_d} = (v_d(1), v_d(2), ..., v_d(TIME'))$;
for all $d \in DAY$ do
  for all $t' \in TIME'$ do
    for $j = 1$ to J_d do
      if $t_{on}^{Rep}(j) \le t' \le t_{off}^{Rep}(j)$ then
        $v_d(t') = 1$;
      else
        $v_d(t') = 0$;
      end if
    end for
  end for
end for
return $\boldsymbol{v_d}$;

3.2.3 *Creation of a User Profile*

A user profile is created for each estimated active interval in each cluster. Information in the EPG is used to create user profiles. A user profile is modeled using the vector space model [2], which is commonly used in the field of natural language processing (NLP). TV programs are also modeled using the vector space model. The vector is called a program-vector. Elements of the vector of a TV program are terms in its EPG, which is calculated using $tf - idf$ [2]. We use Japanese TV programs in our experiment (will be explained later); therefore, texts in the EPG are split into terms using morphological analysis [9]. After acquiring program-vectors for an active interval, the proposed method creates a vector representing its user profile. This vector is called a profile-vector. This vector is created averaging across all vector values in the active interval.

3.3 TV Program Recommendation

This subsection explains how to make a recommendation list to the user at time t on day d. The proposed method acquires active intervals including time t and calculated its profile-vector. It also acquires all EPGs of the TV programs available at time t and creates associated program-vectors. Cosine similarities between the profile-vector and each program-vector are subsequently calculated to recommend the top N TV programs with the greatest similarity to the user.

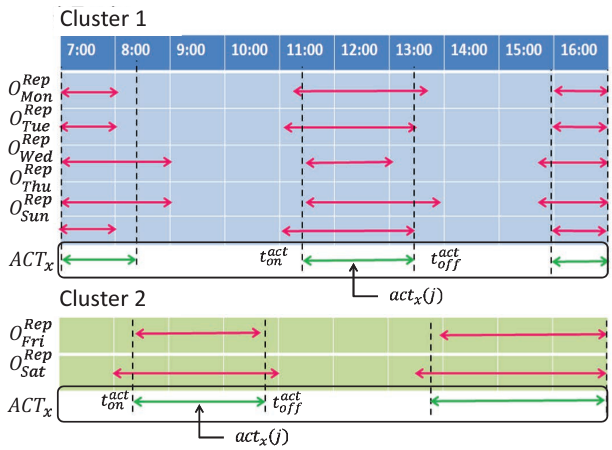


Figure 3: Estimation of active interval.

Table 1: The number of family units.

Members	1	2	3	4	5	6
Family units	0	20	26	34	12	5

4. DATA SET

Viewing logs were collected from Pana-monitors [1] to create a viewing log data set. Data was collected from 97 families over a two week period of times (Sept. 9, 2011 to Sept. 22, 2011) to evaluate the proposed method. Actual EPG data was difficult to collect because of the technical problems; therefore, similar data was collected from online program listings. We refer to this data as the EPG data set.

Pana-monitors were asked to record viewing data for every time they watched TV; specifically, they were asked to record the start time, end time, channel, title of the TV program, its viewers, and the selector of the program watched. The handwritten viewing logs were entered in to a database. A total of 9,442 TV programs were watched. Table 1 shows how many family units (Pana-monitors) exist in the data set for each family size. A family of four members is the most common; Single-person households are not considered as Pana-monitors. The average number of family members is 3.55. Eight-channel data of digital terrestrial broadcast and nine-channel data of digital satellite broadcast were obtained for the EPG data set. The time span of a day was defined as 4:00 to 27:00 in the experiment.

5. EVALUATION

In this section, the effectiveness of the proposed method is verified. Specifically, the following four matters are validated and confirmed: 1) the accuracy of estimating an active interval, 2) the degree of user profile individualization, 3) the accuracy of estimating the program selector, 4) the accuracy

[1] All monitor users are reviewers of products provided by Panasonic Corporation.

of recommended programs. To measure the accuracy of recommended programs, an experiment was conducted. This experiment considered two cases: when all viewing logs are used in the test data and when the viewing logs of habitual TV programs are removed from the test data.

5.1 Evaluation of the Estimation Method for Active Intervals

5.1.1 Evaluation Method

An active interval is the time when the TV is likely to be turned on. The proposed method estimates a daily viewing pattern and generalizes the time that the TV is on to obtain active intervals. Our main objective is to realize TV program recommendations to a dominant user at a specific time. To achieve this goal, an active interval must be estimated correctly. Therefore, we first evaluate the performance of active interval estimation.

The proposed method clusters vectorized time intervals for seven days of the week to estimate a viewing pattern for each day; this method is referred to as $Clustering$. Two additional methods are compared to $Clustering$ to confirm its effectiveness. The first method obtains an active interval using all representative ON/OFF logs of the seven days of the week; in other words, the method does not estimate a viewing pattern for individual days ($AllDay$). The second method obtains an active interval using representative ON/OFF logs on weekdays (Monday through Friday) and weekends (Saturday and Sunday) ($Week$). For $Clustering$, it is necessary to specify the number of clusters in advance. We chose two clusters. As explained in Section 3.2.1, the three generalizing methods used to obtain a representative ON/OFF log are the average ($Average$), median ($Median$), and outlier ($Average^S$) methods. In this experiment, nine combinations of estimation and generalizing methods were compared to estimate daily viewing patterns.

To evaluate the experimental methods, 14-cross validation was used since the viewing log data set was obtained over 14 days. The effectiveness, or precision of estimating a correct time interval, of the nine combinations was confirmed by calculating the similarity between an estimated active interval to the learning data (viewing logs for 13 days) and a set of ON/OFF logs included in the test data (viewing log for a day). The similarity metric used was a cosine similarity between the vector obtained from an estimated active interval and the vector obtained from ON/OFF logs in the test data (correct data). This data was obtained by applying Algorithm 2 to the active interval.

5.1.2 Results and Discussion

Table 2 shows the results of the nine combinations. Each value is an average similarity of 97 families. Among the three generalizing methods, $Median$ achieved the highest similarity. Among the three pattern estimation methods, $Week$ achieved the highest similarity. A t-test was conducted for the three generalizing methods. Significant differences existed between the three methods except for $Average$ and $Average^S$ in the $AllDay$ row ($p < 0.001$). It can be concluded that $Median$ is the best method. A t-test was also conducted for $AllDay$, $Week$, and $Clustering$. $Week$ was significantly different from all other methods except for $AllDay$ in the $Median$ row ($p < 0.001$). Therefore, $Week$ is the best method.

Table 2: Evaluation of estimating active intervals.

	$AllDay$	$Week$	$Clustering$
$Average$	0.4941	0.5180	0.4953
$Median$	0.5676	0.5754	0.5553
$Average^S$	0.4917	0.5307	0.5052

The reason why $Week$ outperformed $Clustering$ is that most of the viewing patterns of families in our data set were classified according to weekday and weekend (Monday - Friday, Saturday - Sunday). If $Clustering$ generated clusters according to weekday and weekend, results identical to $Week$ could be obtained.

5.2 Confirmation of User Profile Individualization

To confirm the degree of individualization of user profiles, we compared our proposed method to the method that creates a single user profile using all the viewing logs (the baseline). This method recommends TV programs using the same user profile for every recommendation time. In this section, we confirm whether a user profile individualized using our proposed method consists of viewing logs of a specific user. If a user profile is created using viewing logs of various users, recommendations should not be tailored to a specific user. On the other hand, if viewing logs used to create a user profile are those of a specific user, recommendations should be specific to the user. Therefore, we investigate the degree of bias of user viewing logs used to create a user profile.

5.2.1 Evaluation Method

As explained in Subsection 3.2.3, a user profile is created using TV programs watched in an active interval estimated for a given recommendation time. Watched TV programs in the dataset contain information about the selector of the program. We identify a user, whose viewing logs are most frequently used, to create a user profile for an active interval. When this ratio is high, a user profile is successfully created for a specific user. This ratio is referred to as the $Occupancy$ and is defined by

$$Occupancy = \frac{Max(n_{u_m})}{n},$$

where, n is the number of viewing logs used to create a user profile. The set of members of family m is denoted by U_m; $u_m \in U_m$ is an individual member of the family. Furthermore, n_{u_m} is the number of viewing logs of the program selector u_m used to create the user profile. $Max(n_{u_m})$ is a function that returns the maximum value of all n_{u_m}; in this case, u_m is redefined as u_{max} (a selector whose viewing logs are used the most frequently to create a user profile). Therefore, the higher the $Occupancy$, the more individualized the user profile.

We used the $Week$ and $Median$ methods to compare the proposed method to the baseline. The $Occupancy$ of each user profile was calculated and its average for each family was obtained. We examined whether or not a user profile created using the proposed method was individualized by comparing the $Occupancy$ averages of the proposed method to those of the baseline.

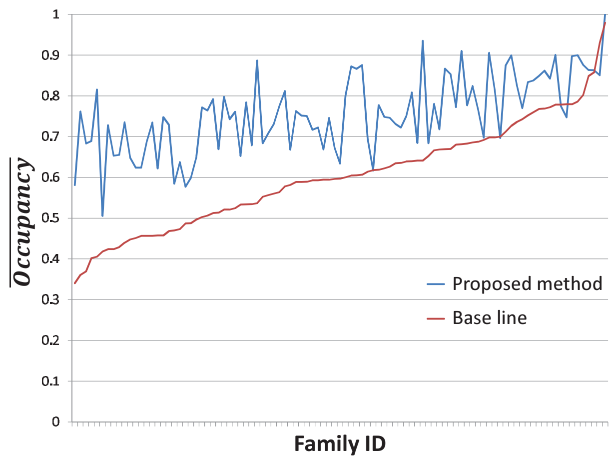


Figure 4: For every active interval in each family, the average Occupancy value is calculated (proposed method). The baseline is the ratio of the most dominant users in the viewing logs of each family.

5.2.2 *Results and Discussion*

Figure 4 shows the *Occupancy* results for individualized user profiles. The values in the figure are the average *Occupancy* values of all families. The value for our proposed method is 0.76 and 0.60 for the baseline. A significant difference existed between the t-test methods ($p < 0.001$). In Figure 4, family IDs are sorted in ascending order of *Occupancy* of the baseline to make the results more comprehensible. Furthermore, observe that there exists a smaller fluctuation (*Occupancy* ranging from 0.6 to 0.9) among families in our proposed method compared to the baseline. This suggests stable results across families.

5.3 Evaluation of Program Selector Estimation Accuracy

The evaluation in the previous section indicated that viewing logs used to create a user profile are biased by logs of a specific user. However, this evaluation does not show whether or not the user profile can realize the recommendation fit for the dominant user based on logs used. Therefore, we investigated method's ability to forecast a dominant user when it recommends TV programs at a specific time.

5.3.1 *Evaluation Method*

To evaluate the estimation of real program selectors, we recommend TV programs that are broadcasted in real-time at the recommendation time using the created profiles. Note that TV programs that are broadcasted in real-time are not used to create user profiles; user profiles are created off line. This profile is used to recommend TV programs that are broadcasted in real-time. We investigated whether the real selector of a recommended TV program matched the dominant user of the user profile. The actual program selector was obtained from the data set. To compare the proposed method to the baseline, we used the *Week* and *Median* methods since they showed the best performance in Subsection 5.1. We used 4:00, 5:00, ..., 27:00 recommendation times and 14-cross validation.

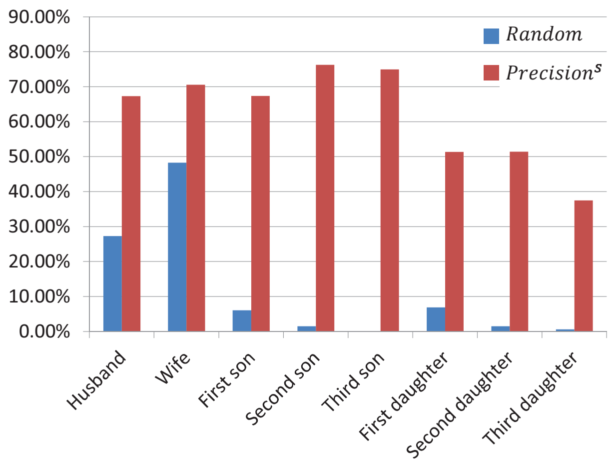


Figure 5: Precision of the program selector

The evaluation metric $Precision^s$ to evaluate the precision of a program selector of recommended TV programs. Specifically,

$$Precision^s = \frac{N_{u^c}}{N},$$

where N is the top N TV programs presented to the user, N_{u^c} is the number of selectors matched to u_{max} in the top N list, and $u_i^c (i = 1, 2, ..., N) \in U_m$ $u_i^c = \phi$ is the selector of the TV program listed in the i-th rank. Note that $u_i^c = \phi$ is possible if no one selected the TV program located in the i-th rank. This evaluation metric was calculated for $N = 1$. Furthermore, we considered a time when evaluation was possible, i.e., when u_i^c exists in the dataset.

5.3.2 *Results and Discussion*

Figure 5 shows the precision results of the program selector ($Precision^s$). The horizontal axis corresponds to a category (relationship) of members in a family. The figure also shows the ratio of each category to all categories of the viewing log data set (denoted as "Random" in the figure). Even though few logs existed for "first son," "second son," "third son," "first daughter," "second daughter," and "third daughter" in our data set, we achieved high $Precision^s$ for these categories. This result suggests that it is possible to create a user profile for family member that infrequently watch TV. The average $Precision^s$ was 67.3% for all categories. Thus, our method recommends TV programs to the dominant user in each time interval when the user frequently watches TV.

5.4 Evaluation of Program Recommendation Accuracy

We validated whether or not an individualized user profile has the ability to recommend TV programs that will be selected by the user. We evaluated the method when all viewing logs in the test data were used and when the viewing logs of habitual programs were removed from the test data.

Table 3: Precision of recommended programs using all viewing logs.

$Precision@N$	N = 1	N = 3	N = 5
baseline	29.3%	18.6%	14.6%
$AllDay$	41.0%	20.1%	14.9%
$Week$	48.9%	22.6%	16.0%
$Clustering$	48.1%	22.3%	15.7%
$7-Days$	55.6%	26.6%	18.0%

Table 4: Precision of recommended programs (habitual viewing logs removed)

$Precision@N$	N = 1	N = 3	N = 5
baseline*	15.8%	13.1%	12.0%
$AllDay*$	29.8%	20.0%	13.6%
$Week*$	34.3%	21.5%	14.9%
$Clustering*$	33.3%	20.3%	14.1%
$7-Days*$	20.9%	14.3%	13.0%

5.4.1 Evaluation Method

We used the $AllDay$, $Week$, $Clustering$, and $7-Days$ methods to estimate daily viewing pattern. $Median$ was used to generalize an active interval. All other conditions of this experiment were identical to those of the previous experiment. To evaluate the precision of recommended programs, we used the following metric:

$$Precision@N = \frac{N_c}{N},$$

where N_c is the number of recommended items. This evaluation was conducted for $N = 1, 3, 5$.

5.4.2 Results and Discussion (full data set used)

Table 3 shows the precision results for recommended TV programs. First, the value of $Precision@N$ was calculated for each recommendation time for each family. Next, for each family, the average $Precision@N$ value across all recommendation times was calculated. Finally, the average value of $Precision@N$ across all families and recommendation times was calculated. Observe from the table that $7-Days$ overcomes the baseline when $N = 1, 3, 5$. An improvement of 26.3% can be seen when $N = 1$, 8.0% when $N = 3$, and 3.4% when $N = 5$. The t-test detected a significant difference ($p < 0.001$). This is a result of the $7-Days$ method using viewing logs from only a day to create a user profile. Moreover, a TV program in the test data is likely to be a habitual one (59.7% of TV programs in our data set are watched for two weeks in a row).

Except for $7-Days$, $Week$ achieved the highest precision. Estimating an active interval with high precision is required to correctly estimate a TV program. A 19.6% improvement can be seen when $N = 1$, 4.0% when $N = 3$, and 1.4% when $N = 5$. Compared to the baseline, t-test detected a significant difference ($p < 0.001$). $AllDay$ and $Clustering$ also overcame the baseline ($p < 0.001$).

The precisions of $Clustering$ and $Week$ were very close; the precision of $AllDay$ was approximately 8% lower than those of $Clustering$ and $Week$ when $N = 1$. This can be attributed to the fact that $AllDay$ uses viewing logs from all days. This prohibits it from distinguishing the dominant users on different days at specific times. On the other hand, $Week$ and $Clustering$, which divide days, separates user profiles when dominant users are different.

These results suggest that the individualization of user profiles in the proposed method allows TV programs to be effectively recommended. In other words, there exists a habitual pattern in the time and content of viewing for each family, which enables the proposed method to recommend TV programs with high precision.

5.4.3 Results and Discussion (habitual TV programs removed)

Although the $7-Days$ method achieved the highest precision in the experiment in the previous section, we suspect that this method weakly generalizes user profile features because it uses viewing logs of the same day. To confirm this, another experiment was conducted in which habitual TV programs (TV programs watched for two weeks in a row) were removed from the data set. Learning data was the same as that used in the previous experiment as well as the evaluation methods used, i.e., baseline, $AllDay$, $Week$, $Clustering$, and $7-Days$. For this experiment, each of these methods was denoted by baseline*, $AllDay*$, $Week*$, $Clustering*$, and $7-Days*$.

The results of this evaluation are shown in Figure 4. As expected, except the $baseline*$, $7-Days*$ obtained the lowest precision. This is because a user profile created by $7-Days*$ consists of a limited number of TV programs watched on the same day. Although $7-Days$ recommends habitual TV programs with high precision, when the habitual TV programs are removed, program recommendations falter as a result of the low ability to generalize user profile features.

We conclude that clustering days of the week according to viewing patterns realizes more accurate recommendations when habitual TV programs are removed.

6. CONCLUSION AND FUTURE WORK

In this paper, a method for recommending TV programs was proposed. The main advantage of this method is that it does not require any user input. Instead, user profiles are created for each time interval that the TV is on. The method was evaluated based on the occupancy of the individualized user profile, the precision of the selector of a TV program, and the precision of recommended TV programs.

Experiments confirmed that user profiles can be generated from specific user logs. Moreover, it was also confirmed that a program selector can be estimated when the user profile is used to recommend TV programs. Experimental results indicate that the proposed method achieves higher precision than other methods that use viewing logs alone, even when habitual TV programs are removed.

EPG terms were used as features in our proposed method. Note that it may be possible to improve the precision treating terms that have the same concept or topic as a similar feature. In the future, we plan to consider Latent Semantic Indexing (LSI) [6] or Latent Dirichlet Allocation (LDA) [5] to reduce dimensions and extract topics.

7. ACKNOWLEDGEMENT

This work was partially supported by the Panasonic Corporation. We would like to give special thanks to Kotaro Sakata and Masaki Yamauchi for their valuable advice and cooperation in collecting viewing log data. We are also deeply grateful to Pana-monitors for filling in our questionnaire about their viewing behavior.

8. REFERENCES

[1] Ali, K., and van Stam, W. Tivo: Making show recommendations using a distributed collaborative filtering architecture. In Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (New York, NY, USA, 2004), KDD '04, ACM, pp. 394–401.

[2] Baeza-Yates, R., and Ribeiro-Neto, B. Modern Information Retrieval: The Concepts and Technology behind Search (2nd Edition) (ACM Press Books), 2 ed. Addison-Wesley Professional, feb 2011.

[3] Bellekens, P., Houben, G.-J., Aroyo, L., Schaap, K., and Kaptein, A. User model elicitation and enrichment for context-sensitive personalization in a multiplatform tv environment. In Proceedings of the Seventh European Conference on European Interactive Television Conference (New York, NY, USA, 2009), EuroITV '09, ACM, pp. 119–128.

[4] Berry, M. J., and Linoff, G. Data Mining Techniques: For Marketing, Sales, and Customer Support. John Wiley & Sons, Inc., New York, NY, USA, 1997.

[5] Blei, D. M., Ng, A. Y., and Jordan, M. I. Latent dirichlet allocation. J. Mach. Learn. Res. 3 (Mar. 2003), 993–1022.

[6] Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., and Harshman, R. Indexing by latent semantic analysis. JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE 41, 6 (1990), 391–407.

[7] Hijikata, Y. Techniques of preference extraction for information recommendation(<special features> exploiting customer's preference : Leading edge of user profiling technique). Information Processing 48, 9 (sep 2007), 957–965.

[8] Ishikawa, E. A comparative table to make sample criterion a for testing rejection. Annual report of the Faculty of Education, University of Iwate 9, 2 (12 1955), 6–15.

[9] Kao, A., and Poteet, S. Natural Language Processing and Text Mining. Springer, 2007.

[10] Ma, Y., Yang, A. Y., Derksen, H., and Fossum, R. Estimation of subspace arrangements with applications in modeling and segmenting mixed data. SIAM Rev. 50, 3 (Aug. 2008), 413–458.

[11] Masthoff, J. Group modeling: Selecting a sequence of television items to suit a group of viewers. User Modeling and User-Adapted Interaction 14, 1 (Feb. 2004), 37–85.

[12] Nakamura, Y., Itou, T., Tezuka, H., Ishihara, T., and Abe, M. Personalized tv-program recommendations based on life log. In Consumer Electronics (ICCE), 2010 Digest of Technical Papers International Conference on (Jan 2010), pp. 143–144.

[13] Seko, S., Yagi, T., Motegi, M., and Muto, S. Group recommendation using feature space representing behavioral tendency and power balance among members. In Proceedings of the Fifth ACM Conference on Recommender Systems (New York, NY, USA, 2011), RecSys '11, ACM, pp. 101–108.

[14] Su, X., and Khoshgoftaar, T. M. A survey of collaborative filtering techniques. Adv. in Artif. Intell. 2009 (Jan. 2009), 4:2–4:2.

[15] Takama, Y., and Muto, Y. Profile generation from tv watching behavior using sentiment analysis. In Proceedings of the 2007 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology - Workshops (Washington, DC, USA, 2007), WI-IATW '07, IEEE Computer Society, pp. 191–194.

[16] Yu, Z., Zhou, X., Hao, Y., and Gu, J. Tv program recommendation for multiple viewers based on user profile merging. User Modeling and User-Adapted Interaction 16, 1 (2006), 63–82.

[17] Zhang, A., Fawaz, N., Ioannidis, S., and Montanari, A. Guess Who Rated This Movie: Identifying Users Through Subspace Clustering. ArXiv e-prints (Aug. 2012).