# An Automatic Personal TV Scheduler based on HMM for Intelligent Broadcasting Services

*Agus Syawal Yudhistira *김문철 **김휘용 **이한규
*한국정보통신대학교 (ICU)
**한국전자통신연구원 (ETRI)
*{agus, mkim}@icu.ac.kr **{hykim5, lhk}@etri.re.kr

# An Automatic Personal TV Scheduler based on HMM for Intelligent Broadcasting Services

*Agus Syawal Yudhistira *Kim, Munchurl **Kim, Hui Yong **Lee, Han Kyu
*Information and Communications University (ICU)
**Electronic and Telecommunications Research Institute (ETRI)

## Abstract

In the future television broadcasting a flood of information from various sources will not always be welcomed by everyone. The need of accessing specific information as required is becoming a necessity. We are interested to make the life of television consumer easier by providing an intelligent television set which can adaptively proposed certain shows to the viewer based on the user historical consumed shows. The TV watching history data consists of TV program titles with their respective genres, channels, watched times and durations, etc. The method proposed is by utilizing Hidden Markov Model (HMM) to model the user preference of kind of genres the viewer will watch based on recorded genres of several weeks time. We take watching schedule from 6 PM to midnight as boundary. The range thus divided into 3 independent time band of 2 hours each resulting in 3 time bands from 6 PM to 8 PM, 8 PM to 10 PM, and lastly 10 PM to midnight. Each time band will be represented by an HMM. From each HMM we can generate a sequence of predicted genre that the user will probably watch during corresponding time-band. Our approach assumes that the user shows a consistent behavior of watching pattern in week to week basis and during the moment of watching TV. To asses the method performance experiment is conducted using real data collected from December 2002 to May 2003. Some user's data are selected and based on that predictions are made. The resulting predictions are then compared with the actual user's history. The experiment shows satisfactory result for user with middle to high consistent behavior level.

## 1. Introduction

Television broadcasting services are entering a new era especially with availability of digital TV, internet, etc. which profoundly affect the TV viewers. Flood of TV programs coming from different sources does not always being welcomed or desirable. Availability of many TV programs at the TV viewer's side entails the difficulty of finding their preferred TV programs. The TV viewers must set aside a considerable amount of their time for searching TV programs and tailoring into their personalized schedules the available published TV program schedules. Worst is the situation when no TV program schedule is available so the TV viewers must find their preferred TV program contents by hopping across the TV channels. The TV viewers can even miss their preferred contents while searching and tailoring TV program schedules for themselves. Our interest is to make it possible the generation of personalized TV program schedules. The personalized TV program scheduler is based on Hidden Markov Model (HMM) by utilizing the TV viewing history data.

The HMM is a well known method to model and predict the processes based on available past behavior information and data. In this paper, our objective is to build a model that predicts the TV watching behaviors in the chronological order of TV genre that a TV viewer is likely to watch. We assume that a TV viewer exhibits a consistent TV watching pattern in his/her TV watching history data. The transition from genre to genre in a day while watching TV can be considered a stochastic process that can be modeled based on the HMM. For the modeling of the chronological order for the watched TV program sequences for each day, we explicitly couple the genre-channel in building our model.

The TV watching history data consists of TV program titles with their respective genres, channels, watched times and durations, etc. To model our automatic personal TV scheduler, the HMM is trained with the TV watching history data for each day. The whole possible TV watching time band is defined from 6:00 P.M. to 12:00 A.M. (midnight)for one day. We assume that the average duration of TV watching time on a specific program is greater than 20 minutes. So the whole time band for one day is segmented into 20 minutes subintervals. The TV watching behavior is observed as follows: (1) it is observed what genres of TV programs a TV viewer has watched every 20 minutes (2) the transition probabilities of the genres are computed every 20 minutes; (3) the TV watching behavior is then regarded as the most probable transition sequence of genres in TV programs; (4) and an estimate sequence of TV channels can be presented to the TV viewer according to the most probable transition sequence of genres.

This paper is organized as follows: Section 2 briefly introduces the HMM for a self contained purpose; Section 3 presents the modeling of a TV personal scheduler by applying the HMM for the usage history data of TV watching; Section 4 shows the experimental results and we conclude our approach in Section 5.

## 2. Hidden Markov Model

HMM is a statistical model of sequential data with double stochastic process. One of the processes is not directly observable and termed as hidden. Observation is conducted trough the other stochastic process that produces a sequence of observable data. This observable stochastic process is a probabilistic function of the hidden stochastic process. Although we say the underlying stochastic process is hidden, this does not mean that the states are totally modeled arbitrarily by means of guess. As in our problem, it is possible for us to observe the channel or genre transition directly and to know how many channels and genres are available. But in some other areas, this information may not be available.

One application that extensively uses HMM is speech recognition. But HMM can also be applied for many kinds of problems such as image analysis, language identification, DNA sequencing, handwriting and text recognition, signal processing, climatology and applied also for many other problems. Defined formally, an HMM $\lambda$ is a 5-tuple of

$$\lambda = (S, V, \Pi, A, B) \qquad (1)$$

where $S = \{s_1, s_2, \ldots, s_N\}$ is a finite set of $N$ states, $V = \{v_1, v_2, \ldots, v_M\}$ is a set of $M$ possible symbols that can be emitted from each states, $\Pi = \{\pi_i\}$ are the initial state probabilities, $A = \{a_{ij}\}$ are the state transition probabilities, $B = \{b_j(k)\}$ are the output or emission probabilities ($B$ is also called as the confusion matrix). In compact and ordinary form, the definition is written as a triplet:

$$\lambda = (\Pi, A, B) \qquad (2)$$

A more detailed of each parameter is as follows:

1. $N$, is the number of states in the model. Thus although the states can be hidden, we are implying that the state is finite.
2. $M$, is the number of distinct observation symbols per state.
3. $\pi_i$ is the probability that the system starts at state $i$ at the beginning.
4. $a_{ij}$ is the probability of going to state $j$ from state $i$.
5. lastly, $b_j(k)$ is the probability of emitting symbol $v_k$ at state $j$.

Of course following constraint must be fulfilled as it is in every Markov model:

$$\sum_{i=1}^{N} \pi_i = 1 \qquad (3)$$

$$\sum_{j=1}^{N} a_{ij} = 1 \quad \text{for} \quad 1 \le i \le N \qquad (4)$$

$$\sum_{k=1}^{M} b_j(k) = 1 \quad \text{for} \quad 1 \le j \le N \qquad (5)$$

These constraints assume that the model is a stationary process and will not vary in time. In our case, two special non-emitting states are added: a starting state $s_{start}$ and an ending state $s_{end}$ in addition to the other ordinary emitting states. These states do not have output probability distributions associated with them but they have transition probabilities. $s_{start}$ is always the first state of the model from which transition to other states begins. Thus, the transition probabilities of this state are the initial state probabilities $\Pi$ itself. $s_{end}$ always comes last toward which the transitions from other states converge. No other transition is possible from $s_{end}$. For more details, readers are recommended to refer to [1][2]. There are 3 problems to solve with HMM:

1. Given the model $\lambda$ and observation sequences, how to compute the probability of the particular output? This is a problem of evaluation.
2. Given the model $\lambda$ and observation sequences, how can we find a sequence of hidden states that best explain the observations? This is a problem of decoding.

3. Given the observation sequences, how can we optimize the model parameters? This is a problem of training.

For training the HMM, we can choose channels or genres as our observation sequences. This training problem is solved by the Baum-Welch algorithm, also known as the forward-backward algorithm. This method consists of two recursion computation: the forward recursion and backward recursion. We use the Baum-Welch algorithm to train our HMM based model [4].

## 3. Modeling a Personal TV Scheduler using HMM

The availability of TV watching history data are necessary to employ the HMM: that is what genre of the program contents the person watched, when they have been watched, how long it was watched and in which TV channel it was broadcasted. In order to implement the HMM to estimate the chronological order of the genre sequences during a day, we choose the genres as the underlying states and the channels as the observations. We assume that TV viewers are allowed to access all the available channels and to watch the program contents in all kinds of genres.

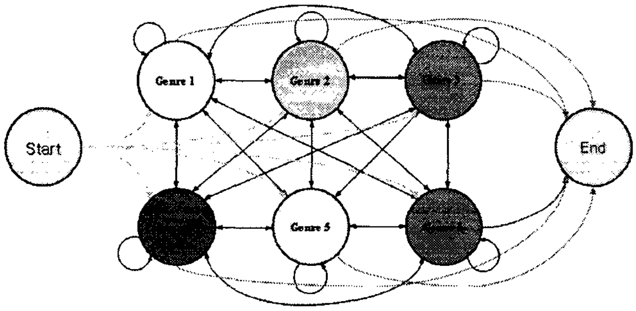


**Fig. 1.** An example of six states (genre) fully connected Markov model. Both *start* and *end* states are non-emitting, while the other six states emitting observation symbols (channels). The number of states is determined by genre types.

TV viewers usually watches TV program contents on some different channels at anytime. This will present us with an ergodic model as shown in Figure 1. Notice that there is no edge connecting the *start* with *end*.

In order for the HMM to be an effective model, the stochastic process should be stationary, not varying in time. Since TV program schedules do significantly vary in time during the course of the days, weeks and months, a significant range of time can be taken and considered invariant. We divide one day into the time bands of 2 hours, each of which bands is represented by one HMM. Since TV programs are usually broadcast in a weekly cycle, the same time bands in the same days for weeks can be considered invariant so the history data of TV watching from the same day during several weeks can be used for training. This is represented graphically in Fig. 2.

As aforementioned, the time band of two hours is segmented into the sub-time bands of 20 minutes long, thus producing six sub-time bands. Each sub-time band represents an observation instance. In other words, we are checking the possibility of changing TV program contents every 20 minutes. Each 20 minutes period is represented by one genre, which is the one that compose most of the 20 minutes time. For instance, if during that 20 minutes time, the viewer is watching *genre1* for 12 minutes and watching *genre2* for the rest 8 minutes, *genre1*is chosen to represent the period. This will give us a coarse estimation on what genre was actually consumed and what genre is expected. Both the genre and channel information of the day is subject to the representations in 20 minutes unit.

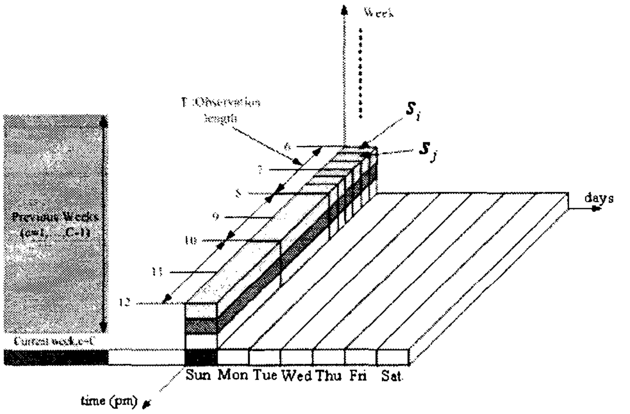


**Fig. 2.** During a week cycle, a television viewer history is divided into days. For each days the data is divided into 2 hours bands, each represented by an HMM.

For training the HMM, a windowing method is employed. We use training data obtained during 8 weeks for which the TV program schedules have maintained unchanged.

### 3.1. Initial HMM Parameters

Rabiner states that although in theory the reestimation equations should give values of the HMM parameters which correspond to a local maximum of the likelihood function, experience has shown that either random (subject to the stochastic and the non zero value constraints) or uniform initial estimates of the $\Pi$ and $A$ parameters is adequate for giving useful reestimates of these parameters in almost all cases [1].

However for the $B$ parameters, good initial estimates are helpful in the discrete symbol case [1]. Given the sequences of genres $G_r=\{g_r(t)\}$ and channels $C_r=\{c_r(t)\}$, $1\leq r \leq R$, $1\leq t \leq T_r$ , where here $R$ is the number of weeks inside a window and $T_r$ is total number of genre sub-time band in week $r$. The component of $B$, $b_j(k)$ is

$$b_j(k) = \frac{\sum_{r=1}^{R}\sum_{t=1}^{T_r} match(c_r(t)=s_j, g_r(t)=v_k)}{\sum_{r=1}^{R} T_r} \tag{6}$$

$$\begin{aligned} match(c_r(t)=s_j, g_r(t)=v_k) &= 1 \quad \text{if } c_r(t)=s_j \wedge g_r(t)=v_k \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{7}$$

## 3.2. Employing the HMM

We employ the HMM for prediction as follows: after the training is completed, a trained HMM is available with new parameters. Following the initial transition $\Pi$ and parameter $A$, as the case of ordinary Markov process, a sequence of predicted genres will be obtained. From the refined parameter $B$, we can deduct the channel for each corresponding genre in the sequence.

But the genre prediction does not always agree with what is actually broadcasted by each TV channel during the day. Adjustment is made according to the prediction of channels and the actual schedule of each TV channel during the day. For instance, for period 7 P.M. to 7:20 P.M, result of prediction is *genre 1* from *channel A*, but according to the day schedule for *channel A*, it should be *genre 2*, then we correct the prediction to be *genre 2* from *channel A*. We termed this adjustment attempt as synchronization.

## 4. Experiment and Results

These experiments utilize the data of TV program viewing history from AC Nielson Korea Research Center, recorded by 2,522 people from December 2002 to May 2003. The history data consists of the following database fields:

| Field Name | Description |
|---|---|
| id | TV viewer's ID |
| date | program broadcasting date |
| dayofweek | a day of the week for program: |
| subscstart_t | beginning time point of watching a program |
| subscend_t | ending time point of watching a program |
| programstart_t | scheduled beginning time of program |
| programend_t | scheduled ending time of program |
| channel | channel of program (6 channels) |
| genre1 | genre of program (8 genres) |

**Table 1.** Experiment source database structure

An attention must be taken regarding the column *genre1*. The column does not provide reliable information on what genres a viewer watches but only provide information of what genres of TV program contents were broadcast before the viewer changes a channel. It is possible though to reconstruct the right schedule table of each channel from the database by aggregating *programstart_t* and *programend_t* of several viewers. We use this resulting table to provide reliable genre information for our experiment performance evaluation.

The six television channels in our experiment include *MBC*, *KBS1*, *KBS2*, *SBS*, *EBS* and *iTV*. Each of the broadcast channels provides the TV program contents in the eight genres: *Education*, *drama & movie*, *news*, *sports*, *children*, *entertainment*, *information* and *others*. Also in the experiments, we define the whole time band for each day from 6 P.M. to midnight only.

Data for the experiments is taken form several people who exhibit their TV watching behaviors with medium to high consistency week by week. The consistency here means that a TV viewer watches a television almost every week for 6 months, and during the day he/she usually watches the television from 6 PM to midnight. It also means that each of the TV viewers exhibits a similar pattern in the chronological order of TV program genres that he/she has watched during the six months.

Preparation for training as follows: First both channel and genre information during 8 weeks periods is filtered. Those viewing record of less than 5 minutes are regarded as hopping and then excluded. The record then is divided into 2 hours bands and 20 minutes sub-time bands. Thus we can have 6 transitions for each 2 hours time band. Since we have 26 weeks recorded data in our database, we can make experiment to predict 18 consecutive weeks. Each week is predicted by using previous 8 weeks recorded data, and the expected result is validated using actual recorded data of the week we try to estimate as shown in Fig. 3.

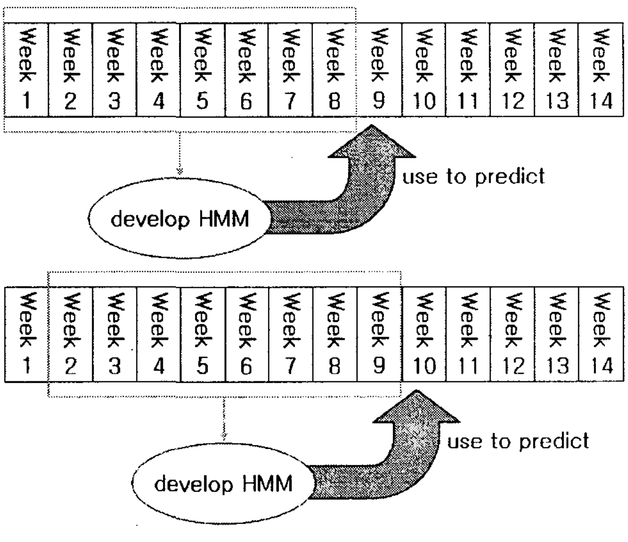


**Fig. 3.** HMM training scheme using 8 weeks previous recorded history

Using 8 weeks data, parameters of the HMM are built and then trained using channel transitions information. Predicted transitions of channel and genre can be obtained from the trained model.

The evaluation of the trained HMM is shown in Figure 4. The evaluation is made as follows: for each sub-time band of 20 minutes where the prediction match with actual choice the viewer made as recorded in the database, a score is given one. Those that do not match are given zero score. This is done for the sequence of predicted channels and genres, respectively. Consistency of the person is computed with the same method by comparing two consecutive weeks. Given a set of genre transition $G_W = \{g_w(t)\}$, a set of prediction of genre transition $G'_W = \{g'_w(t)\}$, $1 \le w \le W$ and $1 \le t \le T_w$, where here $W$ is the number of experiment weeks we predict and $T_W$ is total number of genre sub-time band in week $w$. Expected genre transition accuracy is

$$\text{genre prediction accuracy} = \frac{\sum_{w=1}^{W}\sum_{t=1}^{T_w} genre\_match(g'_w(t), g_w(t))}{\sum_{w=1}^{W} T_w} \tag{8}$$

$$genre\_match(g'_w(t), g_w(t)) = 1 \quad \text{if } g'_w(t) = g_w(t) \\ = 0 \quad \text{otherwise} \tag{9}$$

The genre consistency is counted as

$$\text{genre consistency} = \frac{\sum_{w=2}^{W}\sum_{t=1}^{T_w} genre\_match(g_{w-1}(t), g_w(t))}{\sum_{w=1}^{W} T_w} \tag{10}$$

$$genre\_match(g_{w-1}(t), g_w(t)) = 1 \quad \text{if } g_{w-1}(t) = g_w(t) \\ = 0 \quad \text{otherwise} \tag{11}$$

And given a set of channel transition $C_W = \{c_w(t)\}$, a set of prediction of channel transition $C'_W = \{c'_w(t)\}$, $1 \le w \le W$ and $1 \le t \le T_w$, where here $W$ is the number of weeks we predict and $T_W$ is total number of channel sub-time band in week $w$. Expected channel transition accuracy is

$$\text{channel prediction accuracy} = \frac{\sum_{w=1}^{W}\sum_{t=1}^{T_w} channel\_match(c'_w(t), c_w(t))}{\sum_{w=1}^{W} T_w} \tag{12}$$

$$channel\_match(c'_w(t), c_w(t)) = 1 \quad \text{if } c'_w(t) = c_w(t) \\ = 0 \quad \text{otherwise} \tag{13}$$

And channel consistency is counted as

$$\text{channel consistency} = \frac{\sum_{w=2}^{W}\sum_{t=1}^{T_w} channel\_match(c_{w-1}(t), c_w(t))}{\sum_{w=1}^{W} T_w} \tag{14}$$

$$channel\_match(c_{w-1}(t), c_w(t)) = 1 \quad \text{if } c_{w-1}(t) = c_w(t) \\ = 0 \quad \text{otherwise} \tag{15}$$

Selection of specimen is based on the consistency of the viewer, neglecting the viewer's profile.

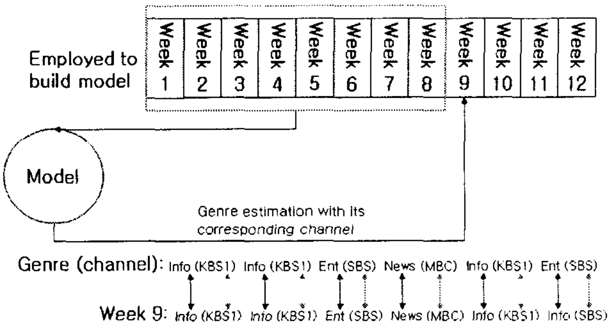


**Fig. 4.** Evaluating expected transition of genres (and channels) of week 9 with actual recorded data of that week.

Table 2 and 3 represent the experimental results of specimens under two different scenarios. Table 2 and 3 represent the prediction accuracies with the genre as state and with the channel as state, respectively. Even with the data with high consistency, Table 3 shows lower accuracy results compared to that in Table 2. Especially, the channel prediction is likely to fail along with genre.

Under the heading *estimation*, there are three columns. First second column are result of prediction without adjustment. The last column, *sync. genre*, shows accuracy of genre prediction with synchronization as stated in section 3.2. This cannot be done in the scenario which results in Table 3. User ID 115434206 (Wednesday, second model) shows a little anomaly, which can be explained. During 2 hours period the person watched extensively one kind of genre provided from several channels. The high accuracy of genre estimation (without synchronization) shows that the prediction stays in one kind of genre. Yet, for each state, only one channel will be assigned highest probability. The transition of channels is not adequately reflected in this case and thus lowers accuracy result. This can be an indication that 2 hours band is not stationary enough, but overall result is quite good for coarse deduction and content advising. Those that less consistent in TV consumption showed as expected low prediction accuracy. Reason why the second scenario does not provide good result is because the confusion matrix (genre) only reflects distribution of genre broadcasted by the channel. Thus, if the channel failed to produce good prediction the genre will also fail. This is different from the first scenario where the genre probability is summarized from all possible channels available.

| User ID | Day of Week | HMM 1 (6 PM - 8 PM) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Viewing Consistency | | estimation | | | |
| | | genre | channel | genre | channel | sync genre | |
| | | | | | | genre | channel |
| 115434206 | Tuesday | 85.19% | 96.30% | 67.59% | 98.15% | 98.15% | 98.15% |
| 115434206 | Wednesday | 84.11% | 91.59% | 47.66% | 96.26% | 96.26% | 96.26% |
| 125444502 | Wednesday | 76.64% | 62.62% | 17.76% | 57.94% | 67.29% | 57.94% |
| 13020602 | Wednesday | 39.39% | 37.37% | 19.19% | 36.36% | 44.44% | 36.36% |
| 113431102 | Friday | 71.59% | 71.59% | 70.45% | 80.68% | 87.50% | 80.68% |
| | | HMM 2 (8 PM - 10 PM) | | | | | |
| 115434206 | Tuesday | 90.74% | 94.44% | 61.11% | 81.48% | 82.41% | 81.48% |
| 115434206 | Wednesday | 80.00% | 82.86% | 89.52% | 51.43% | 54.29% | 51.43% |
| 125444502 | Wednesday | 77.36% | 72.64% | 51.89% | 64.15% | 82.08% | 64.15% |
| 13020602 | Wednesday | 37.25% | 34.31% | 18.63% | 36.27% | 42.16% | 36.27% |
| 113431102 | Friday | 89.50% | 87.62% | 35.24% | 94.29% | 94.29% | 94.29% |
| | | HMM 3 (11 PM - 12 PM) | | | | | |
| 115434206 | Tuesday | 82.28% | 79.75% | 62.03% | 86.08% | 88.61% | 86.08% |
| 115434206 | Wednesday | 62.86% | 44.29% | 62.86% | 65.71% | 75.71% | 65.71% |
| 125444502 | Wednesday | 73.17% | 67.07% | 32.93% | 51.22% | 67.07% | 51.22% |
| 13020602 | Wednesday | 72.12% | 63.46% | 57.69% | 70.19% | 77.88% | 70.19% |
| 113431102 | Friday | 94.44% | 90.74% | 47.22% | 95.37% | 97.22% | 95.37% |

**Table 2.** Prediction accuracy: *Genre* as state and *channel* as observation.

| User ID | Day of week | HMM 1 (6 PM - 8 PM) | | | |
|---|---|---|---|---|---|
| | | viewing consistency | | Estimation | |
| | | genre | channel | channel(state) | genre(observ) |
| 115434206 | Tuesday | 85.19% | 96.30% | 65.42% | 49.53% |
| 115434206 | Wednesday | 84.11% | 91.59% | 92.59% | 75.00% |
| 125444502 | Wednesday | 76.64% | 62.62% | 22.43% | 56.07% |
| 13020602 | Wednesday | 39.39% | 37.37% | 31.31% | 17.17% |
| 113431102 | Friday | 71.59% | 71.59% | 77.27% | 72.73% |
| | | HMM 2 (8 PM - 10 PM) | | | |
| 115434206 | Tuesday | 90.74% | 94.44% | 50.00% | 87.25% |
| 115434206 | Wednesday | 80% | 82.86% | 38.89% | 60.19% |
| 125444502 | Wednesday | 77.36% | 72.64% | 50.00% | 61.32% |
| 13020602 | Wednesday | 37.25% | 34.31% | 18.63% | 24.51% |
| 113431102 | Friday | 89.50% | 87.62% | 71.43% | 54.29% |
| | | HMM 3 (10 PM - 12 PM) | | | |
| 115434206 | Tuesday | 82.28% | 79.75% | 51.56% | 67.19% |
| 115434206 | Wednesday | 62.86% | 44.29% | 82.28% | 68.35% |
| 125444502 | Wednesday | 73.17% | 67.07% | 32.14% | 45.24% |
| 13020602 | Wednesday | 72.12% | 63.46% | 68.27% | 58.65% |
| 113431102 | Friday | 94.44% | 90.74% | 88.89% | 55.56% |

**Table 3.** Prediction accuracy: *Channel* as state and *genre* as observation.

## 5. Conclusion

In this paper, we propose an automatic personal TV scheduler that is modeled using HMM. The usage history data of TV watching has been used to train and to test the proposed personal TV scheduler. The double stochastic property of HMM is exploited to synchronize the channel preference with the expected genre preference. It provides satisfactory accuracy to predict TV viewer's preference for middle to high consistent TV consumption.

## Reference


1. Rabiner, L.R.: ATutorial on Hidden Markov Models and Selected Application in Speech Recognition. Proc. IEEE 77 (1989) 257-285.
2. Zhai, Cheng Xiang: A Brief Note on the Hidden Markov Models (HMMs). Department of Computer Science University of Illinois at Urbana-Champaign (2003).
3. Stamp, Mark: A Revealing Introduction to Hidden Markov Models. Department of Computer Science San Jose State University (2004).
4. Young, S., Evermann, G., Gales, M., et al.: The HTK Book. Cambridge University Engineering Department (2005) 6-8, 128-130.