Exploring Interpersonal Influence by Tracking User Dynamic Interactions

Chengyao Chen and Wenjie Li, Hong Kong Polytechnic University
Dehong Gao, Alibaba Group
Yuexian Hou, Tianjin University

A novel temporal influence model can learn users’ opinion behaviors by exploring how influence emerges in social media communications.

Social media networks such as Twitter allow people to communicate and share information, opinions, and experiences. Through these interactions, influence on thoughts and behaviors emerges. Today, opinion influence has attracted tremendous interest from marketers with the aim of increasing brand awareness and sales.

The research on measuring users’ global influence and their different roles over the whole network provides a good starting point to developing influence-driven business strategies. However, when limited by budgets, we want to locate a specific number of users whose influence reaches the greatest possible range on the network, a research problem known as influence maximization. To solve this problem, understanding interpersonal influence is the key and the focus of this article.

In the past literature, researchers have proposed diffusion-based influence models such as Independent Cascade (IC) and General Threshold (GT). These models share the same assumption: that influence represents the ability to successfully activate state change. Because these models totally ignore the effects of influence during a long period of dynamic interactions, they aren’t sufficient for capturing how people form and change opinions about a specific product or topic of interest. We aim to model interpersonal influence by tracking and correlating users’ opinion behaviors through their dynamic interactions.

Coupled hidden Markov models (HMMs) have been successfully applied to complex action recognition by modeling the interacting processes of two hands or legs. The observed Markov chains are also employed to model the speaking/silent action behaviors of participants in small-scale laboratory experiments. Similar to diffusion-based approaches, the interaction-based influence model is a binary model that’s concerned with two states in discrete times, focusing on the estimation of state transition probabilities.
Related Work in Influence Modeling

Although the influence maximization problem is a popular research topic, the interpersonal influence between two users that triggers the diffusion of information is totally ignored. Saito and colleagues developed the EM algorithm to learn diffusion probability based on information diffusion using the Independent Cascade (IC) model. Similarly, Goyal and colleagues adopted the General Threshold (GT) model to develop a variety of probabilistic models measuring the diffusion probability. Furthermore, Miller and colleagues explored how sentiment diffuses on social media based on the same assumption that influence is only effective when a user is activated. However, the instant influence assumption totally ignores the effects of influence during long periods of dynamic interactions, and one-time activation instances aren’t enough to learn influence.

To recognize complex action, coupled hidden Markov models (HMMs) have been successfully applied to action recognition by modeling the interacting processes of two hands or legs. Specifically, they model the activity of each hand as a Markov chain and try to understand what people are doing via interactions between two hands. Following this idea, a simplified coupled-HMM influence model was first theoretically studied by Asavathiratham to understand the behaviors of large number of interacting components in a complex network, such as a communication network or a transportation system. Basu and colleagues employed a similar idea to learn speaking/silent action behaviors of participants in small-scale laboratory experiments. All these approaches assume that the actions of all participants change simultaneously, and thus can’t handle content-based behaviors such as opinions, which differ a lot among users.

With the explosion of social media, understanding sentiment from opinion-rich data becomes more and more important. Hu and colleagues proposed using emotional signals for unsupervised sentiment classification. Social relations have also been used to enhance sentiment classification. All these methods try to detect sentiment from social text, but they fail to explore the reason why users express opinions. In our work, taking posting sentiment as a premise, we try to learn social influence among users and incorporate interpersonal influence to predict users’ future opinions.

References


Data Collection

To explore interpersonal influence, we start with four widely discussed electronic products—iPhone, Samsung Galaxy, Xbox, and PlayStation. For each product, we construct a dataset by first collecting all tweets containing the topical word (such as “iPhone”) from 1 August 2014 to 31 October 2014 (three months) with the Twitter streaming API (dev.twitter.com/streaming/overview) and then identifying each tweet’s opinion. Among all the users, those who post more than 10 tweets in total constitute the active user set. The “following” relationships of active users are collected via the Twitter REST API; we use them to construct independent friend networks, from which we then select those with more than 1,000 users for analysis.
Table 1 shows the dataset’s statistics, including the numbers of users, relationships, and user networks. All tweets are associated with their posting times (in minutes).

**Opinion Analysis**

A tweet’s opinion polarity is determined based on a publicly available sentiment lexicon called AFINN (www2.compute.dtu.dk/~faan/data). It’s manually constructed for sentiment analysis on microblogs, with each word receiving an integer value between –5 and +5 denoting its sentiment strength. For each tweet, we obtain its overall opinion score by summing the sentiment scores of the words included—such as whether the tweet contains negation words such as “no” and “nothing” or adversative conjunctions such as “however” and “but.” Opinion polarity is then derived from the overall opinion score. A positive score denotes a positive opinion, a negative score denotes a negative opinion, and a zero score denotes a neutral opinion.

**Shuffle Test**

Before learning influence from the interactions of Twitter users, we need to prove the existence of influence on Twitter itself. Anagnostopoulos and colleagues used the shuffle test to verify the existence of social influence on a social network. Motivated by this idea, we assume that if influence isn’t a likely source of opinion correlation, then the timing of friends’ opinions shouldn’t matter, so reshuffling the time stamps of posted opinions shouldn’t significantly change the amount of correlation. The following shuffle test helps verify the correlation between changes in someone’s opinion and a friend’s based on this assumption.

<table>
<thead>
<tr>
<th>Topic</th>
<th>No. users</th>
<th>No. relationships</th>
<th>No. networks (&gt; 1,000 users)</th>
</tr>
</thead>
<tbody>
<tr>
<td>iPhone</td>
<td>174,699</td>
<td>2,585,184</td>
<td>30</td>
</tr>
<tr>
<td>Samsung Galaxy</td>
<td>17,812</td>
<td>192,406</td>
<td>5</td>
</tr>
<tr>
<td>Xbox</td>
<td>48,314</td>
<td>696,708</td>
<td>14</td>
</tr>
<tr>
<td>PlayStation</td>
<td>45,001</td>
<td>1,039,328</td>
<td>7</td>
</tr>
</tbody>
</table>

When user u posts an opinion o that differs from his or her typical opinion, we call it an activation instance, and the activation time is t. All u’s friends who post before t and after u’s previous tweet provide the background information for the activation; we call them u’s context friends. Those who post the same opinion o as u are called activation friends, and they could be the source for why u changes opinions. Here, we use the percentage of activation friends out of all context friends to represent opinion context before activation.

We then construct the shuffled dataset by permuting the sequence of tweets for each user randomly, which changes the opinion context for each activation instance. Figure 1 plots correlations between activation and opinion context based on the biggest network in the “iPhone” dataset with 44,889 users, 685,526 relationships, and 2,333,804 tweets. If influence doesn’t exist in the network, correlations on the shuffled dataset and the original dataset should follow the same or a very similar curve. However, as observed in Figure 1, the conditional distribution function in the shuffled dataset stands significantly on the left side of the one in the original dataset; in the original dataset, more instances of opinion changes are correlated with friends’ opinions. This verifies that opinion changes have stronger correlations with friends’ opinions in the original dataset compared with the shuffled dataset. It also proves the existence of influence as expected.

Note that similar results are also found in the other datasets—we omit displaying them here due to space constraints.

**Our Method**

To track users’ opinion behaviors under the social influence, we propose a temporal influence model (TIM) that integrates opinion-preserving and influence-decaying assumptions.
Problem Formulation
Let $G = (U, R)$ be a directed graph. $U$ is a set of $|U| = N$ users, and $R = \{u, u'\} u \in U, u' \in U$ denotes relationships among users, where $(u, u') \in R$ means user $u$ follows user $u'$. $S^{(u)} = [s^{(u)}_1, s^{(u)}_2, ..., s^{(u)}_{|S^{(u)}|}]$ represents $u$’s opinion sequence on a specific topic, and $S^{(u)}_t$ is $u$’s opinion at time $t(S^{(u)}_t)$. $S^{(u)}_t \in \Omega$, where $\Omega = \{1, 2, ..., |\Omega|\}$ is the set of possible opinion sentiments. In our work, we assume that users have three types of opinion sentiments, where $\Omega = \{1, 2, 3\}$ corresponds to negative, neutral, and positive sentiments, respectively. The opinion sequences from different users constitute $S = \{S^{(u)} | u \in U\}$.

Given $U$, $R$, and $S$, our objective is to learn the influence strength, denoted as $a_{uw}$, representing the power of influence that $u'$ exerts on $u$, and the influence effects of $u'$ on $u$, denoted as the state transition probability matrix $\varphi^{(u,u')} = \begin{bmatrix} \varphi_{m,n}^{(u,u')} \end{bmatrix}_{3 \times 3}$, where $m, n \in \Omega$. $\varphi^{(u,u)}$ means the state transition under the self-influence.

Temporal Influence Model
From $R$, we can derive a set of influencing users for $u$, represented by $IF_u = \{u' \mid (u, u') \in R\}$, where $|IF_u|$ is the size of $IF_u$. Following tradition, we assume the independence of the influences from each $u' \in IF_u$ on $u$ and the integrated influence as the linear combination of all interpersonal influences from $IF_u$. Thus, the probability of user $u$’s opinion could be formulated as

$$P(S^{(u)} | u, IF^{(u)}_u, \ldots, IF^{(u)}_{|U|}) = \sum_{u'} a_{uw} P(S^{(u')} | IF^{(u')}_{u'}) + a_{uu} P(S^{(u)} | u),$$

(1)

where $u' \in IF_u, a_{uw}$ is the power of influence that $u'$ exerts on $u$, and $a_{uu}$ represents $u$’s degree of opinion preservation. We then construct the influence vector $a_{uw} = \begin{bmatrix} a_{uw}, a_{uw}(IF^{(u)}_1), \ldots, a_{uw}(IF^{(u)}_{|U|}) \end{bmatrix}$ to represent the degree of opinion preservation and interpersonal influence strength together; $a_{uw}$ could be referred to as the probabilities of which chain (including friends and his or her own) $u$ chooses to determine state. The sum of $a_{uw}$ is therefore 1. If $a_{uu}$ is up to 1, it means that $u$ is difficult to be influenced by others.

Furthermore, we put forward the following additional hypotheses.

Opinion-preserving assumption. People tend to maintain their own personal opinions before they’re convinced and influenced by others. We assume self-influence is the long-term preference distributed over the three possible sentiments and simplify $\varphi^{(u,u)}$ as $\varphi^{(u,u)} = \begin{bmatrix} \theta_{u}^{(u)} \end{bmatrix}_{3 \times 1}$ and $\Sigma \theta_{u}^{(u)} = 1$, where $s \in \Omega$.

Influence-decaying assumption. During communication, the information $u$ receives since the time of his or her previous tweet has influence on his or her current opinion. Therefore, there might exist more than one opinion from $u$’s friend $u'$ that influences $u$. That said, the effects of multiple opinions at different times aren’t the same. Instead, their effects gradually decrease over time with decay behaviors. We use the exponential function to model this situation. Then, the aggregated influence from $u'$ can be formulated by weighting the importance of opinions received at different times:

$$P(S^{(u)} | u', w_u^{(u,u')}) = \sum_{i=1}^{N} \frac{1}{e^{-\lambda (t(S^{(u)}_i))}} \frac{1}{e^{-\lambda (t(S^{(u')}_i))}}$$

(2)

where $w_u^{(u,u')} = \frac{1}{e^{-\lambda (t(S^{(u)}_i))}} \frac{1}{e^{-\lambda (t(S^{(u')}_i))}}$, is a normalized term that represents the importance of posts at different times, and $j$ satisfies $t(S^{(u)}_j) < t(S^{(u')}_j) < t(S^{(u)})$.

To this end, we formulate the TIM by combining the above-mentioned assumptions (see Figure 2). The likelihood function is written as

$$P(S) = \prod_{u} P(S_{u}^{(u)})$$

(3)

where $u \in U, u' \in IF_u, w_u^{(u,u')} = \frac{1}{e^{-\lambda (t(S^{(u)}_i))}} \frac{1}{e^{-\lambda (t(S^{(u')}_i))}}$.
The TIM is characterized by \((\Theta, \Phi, A)\), where \(\Theta\) is the personal opinion preference vector, \(\Phi\) is the state transition probability matrix, and \(A\) is the influence strength vector. Our objective is to maximize the likelihood function \(P(S)\) by learning these three parameters. To learn \(\Theta\), we follow the opinion-preserving assumption and estimate \(\Theta\) by the proportion of each opinion sentiment type in the entire opinion sequence directly. In \(\Phi\), each state transition probability matrix between a pair of users is independent of other state transition probability matrices. Meanwhile, for users \(u\) and \(u'\), where \(u\) follows \(u'\), the conditional probability matrix \(\phi^{(u,u')}\) should satisfy \(\sum_{s} \phi_{s,s'}^{(u,u')} = 1\) for \(s \in O\). We infer \(\phi^{(u,u')}\) by using the maximum likelihood estimation (MLE) for Markov chains, which also counts the time-decayed frequency of each state transition type from the two connected chains of \(u\) and \(u'\).

Given \(\Theta\) and \(\Phi\), our objective becomes to maximize \(P(S)\) by inferring parameter \(A\). We rewrite the likelihood function to log likelihood and remove the parts not related to \(A\), observing that the influence vector \(\alpha_u\) of a user is independent of others in the log likelihood function. We thus update \(\alpha_u\) for each user individually. We simplify the objective function by separating it into an individual log likelihood function for each user, as in Equation 4:

\[
\alpha_{u*} = \arg\max_{\alpha_u} \sum_{t=1}^{T} \log \left( \sum_{u' \in U} \alpha_{u'} \sum_{(s_1, s_2)} \phi_{s_1, s_2}^{(u,u')} \phi_{s_2, s_3}^{(u',u''')} \alpha_{u''} \phi_{s_3, s_4}^{(u'',u''')} \right)
\]

\[
\text{subject to } \sum_{u' \in U} \alpha_{u'} + \alpha_u = 1, \alpha_u > 0, \text{ for } u' \in U^N
\]

(4)

By using Jensen’s inequality, this per-chain likelihood function can be proved concave in \(\alpha_u\). To learn the optimal solutions of \(\alpha_u\), we remove the equality constraints by representing \(\alpha_u\) as \(1 - \sum u' \alpha_{u'}\), where \(u' \in IF_u\). All the constraints now become the inequality constraints:

\[
\alpha_{u'} > 0, \text{ and } 1 - \sum u' \alpha_{u'} > 0.
\]

We approximate the objective function by adding the constraints with a logarithm barrier function in Equation 5, where \(\beta_1, \beta_2\), represent approximating accuracy:

\[
\max_{\alpha_u} F = \sum_{t=1}^{T} \log \left( \sum_{u' \in U} \alpha_{u'} \sum_{(s_1, s_2)} \phi_{s_1, s_2}^{(u,u')} \phi_{s_2, s_3}^{(u',u''')} \alpha_{u''} \phi_{s_3, s_4}^{(u'',u''')} \right) + \left(1 - \sum_{u' \in U} \alpha_{u'}\right) \log\left(1 - \sum_{u' \in U} \alpha_{u'}\right) + \beta_1 \log(1 - \sum u' \alpha_{u'}) + \beta_2 \sum u \log(\alpha_u).
\]

To make the approximate solution close to the original solution, we use the barrier method to iteratively solve the unconstrained optimization problem in Equation 5 by decreasing the values of \(\beta_1\) and \(\beta_2\). In each iteration, we apply the gradient descent method to find the optimal solution. The starting point for the current iteration is the optimized point found in the last iteration. For each iteration, the derivative \((\partial F / \partial \alpha_{u*})\) of objective \(F\) with respect to \(\alpha_{u*}\) is calculated, and \(\alpha_{u*}\) is updated using the following rules until the solution converges. \(\gamma\) is the learning rate:

\[
\begin{align*}
\alpha_{u*} &\leftarrow \alpha_{u*} + \gamma \frac{\partial F}{\partial \alpha_{u*}} \\
\end{align*}
\]

(6)

For the \(k\)th element \(\alpha_{uk}\) of the vector \(\alpha_u\), the derivative is

\[
\frac{\partial F}{\partial \alpha_{uk}} = \sum_{t=1}^{T} \log \left( \sum_{u' \in U} \alpha_{u'} \sum_{(s_1, s_2)} \phi_{s_1, s_2}^{(u,u')} \phi_{s_2, s_3}^{(u',u''')} \alpha_{u''} \phi_{s_3, s_4}^{(u'',u''')} \right) - \beta_1 \log(1 - \sum u' \alpha_{u'}) - \beta_2 \sum u \log(\alpha_u)
\]

(7)

Based on these learning rules for the three parameters, the learning steps could be implemented independently for each user. Thus, the implementation of the learning algorithm could be faster by employing multi-parallel tasks or scaling up to distributed systems.

Finally, we perform a time complexity analysis for the learning algorithm. Because the parameters \(\phi^{(u,u')}\), \(\Theta^{(u)}\), and \(\alpha_u\) can be learned for each user \(u\), we perform the analysis for each user \(u\) individually. Leaning the parameter \(\Theta^{(u)}\), the time complexity is \(O(|S^{(u)}| |IF_u|)\). We then analyze the upper bound of the complexity on the barrier method to learn parameter \(\alpha_u\).

The number of Newton steps per outer iteration is \(O(\log(|IF_u|))\), and the number of outer steps required is \(O(|IF_u|)\). After multiplying the cost within one Newton iteration, we get the upper bound for learning parameter \(\alpha_u\), which is \(O(|IF_u|^2 \log(|IF_u|) |S^{(u)}|)\). Putting it together requires \(O(|IF_u|^2 \log(|IF_u|) |S^{(u)}|)\) of time complexity to solve the objective function in Equation 3.

**Experiments and Discussion**

We evaluate our proposed model on the opinion prediction task. Experimental results demonstrate the effectiveness of our proposed model, especially on predicting future opinions for those who change their minds frequently.

**Experiment Setup**

Because the value of interpersonal influence is difficult to verify directly,
we use the opinion prediction task to help us evaluate it extrinsically. For each user $u$, previous $|S^{(u)}| - 1$ opinions and friends’ opinions are selected into the training dataset, and his or her last opinion $S^{(u)}_{u(t)}$ and relevant influencing opinions from friends are kept in the test dataset. We compare our proposed model (TIM) with the three types of methods, specifically, individual-based, influence-based, and unified methods incorporating personal opinion and social influence together:

- **Majority** predicts $u$’s opinion based on the most frequent opinion sentiment observed in past posts. It can also be regarded as personal preference-based prediction.
- **FriOpinion** predicts the $i$th opinion of $u$ from the most frequent opinion from $u$’s friends that post during time $t \left( S^{(u)}_{u(t) - 1} \right) - t \left( S^{(u)}_{u(t)} \right)$. It can
- **IMC** (independent Markov chain) models the opinion behaviors of each user independently as a single observed Markov chain following the first-order Markov assumption.

Table 2. Performance evaluation.

<table>
<thead>
<tr>
<th>Method</th>
<th>Influence-based method</th>
<th>Individual-based method</th>
<th>Unified method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic</td>
<td>FriOpinion</td>
<td>TIM_Fri</td>
<td>Majority</td>
</tr>
<tr>
<td>iPhone</td>
<td>0.2808</td>
<td>0.3024</td>
<td>0.5579</td>
</tr>
<tr>
<td>Samsung Galaxy</td>
<td>0.3691</td>
<td>0.3954</td>
<td>0.6697</td>
</tr>
<tr>
<td>Xbox</td>
<td>0.2915</td>
<td>0.3083</td>
<td>0.5257</td>
</tr>
<tr>
<td>PlayStation</td>
<td>0.3072</td>
<td>0.3708</td>
<td>0.5715</td>
</tr>
</tbody>
</table>

Figure 3. Performances on user groups with different contradiction levels.
also be regarded as voting-based prediction.

- **TIM_Fri** is a simplified version of TIM that considers only friends’ influence but not the opinion-preserving assumption.
- **SVM** (support vector machine) classifies the opinion sentiment of each user into three types, that is, positive, neutral, and negative. At each posting time $t(S^{(u)}_j)$, we construct a training instance with the features, including the previous opinions of $u$ and $u$’s friends, and attach $u$’s opinion sentiment $S^{(u)}_j$ as the label. To be consistent with TIM’s assumptions, the previous opinion of $u$ is his or her major opinion sentiment, and the previous opinions of $u$’s friends are calculated based on the linear combination of the opinion sentiment posted during $t(S^{(u)}_j) - t(S^{(u)}_{j-1})$ with time decay taken into the account.
- **TIM_S1** is a simplified TIM that assumes each user has the same influence on all followers. We employ the PageRank algorithm to learn the influence strength for each $u$. The normalized influence power of $u$ and $u$’s influencers $u'$ are defined as the influence strength $\alpha_{uu'}$ and $\alpha_{uu''}$.
- **TIM_S2** is another simplified TIM that assumes $u$ is equally influenced by all friends, that is, $\alpha_{uu'} = \alpha_{uu''} = \frac{1}{|IF_u|+1}$ for $u' \in IF_u$.

We experimentally set the time decay rate $\lambda$ as 1 when used in TIM, TIM_S1, TIM_S2, TIM_Fri, and SVM.

**Performance Evaluation**

We use the accuracy of opinion prediction to evaluate TIM’s effectiveness:

$$\text{accuracy} = \frac{\text{no. correctly predicted users}}{\text{no. all users in the network}}$$

Table 2 shows the average accuracy over all the networks. Over the four topics, TIM consistently achieves better performance than all baseline methods. We also perform the pairwise $t$-test between TIM and each baseline method. All $p$ values are smaller than 0.05, which means the improvements are significant. In addition, we also observe the following findings.

First, comparing the two individual-based methods, we notice that Majoriy performs much better than IMC. The assumption used in Majority is the same as the opinion-preserving assumption used in our proposed model; its better performance indicates the opinion-preserving assumption is reasonable to model how a user maintains his or her personal opinions.

Second, of the two influence-based methods, TIM_Fri is better than Fri-Opinion. The linear combination is a better representation for the aggregated influence from friends than the voting scheme, which also supports the linear influence assumption used in TIM.

Third, although SVM encodes all the assumptions made in TIM, including the opinion-preserving assumption and influence-decaying assumption, its performance is much worse than TIM and even worse than the two simplified versions of TIM. This is because SVM doesn’t have the ability to quantitatively measure the power of influence, especially how it affects changes of opinion, which is important for predicting user opinions.

Finally, compared with Majority and TIM_Fri, the better performance of TIM proves that the factors of personal preference and social influence are both necessary for predicting opinions. Comparing the two factors, personal preference is more dominant in decision making.

**Performances on Users with Different Degrees of Being Influenced**

The Majority method’s good performance seems to suggest that some users tend to preserve their opinions consistently, which makes it difficult to observe the interpersonal influence on them. To further explore the effectiveness of learned influence, we look at those users who have different degrees of being influenced.

Intuitively, users who change their opinions frequently are more likely to be influenced than those who stick to one opinion consistently. Opinion diversity is therefore an indicator of the degree of being influenced. It can be measured by combining the mean $\mu$ and the variance $\sigma^2$ of all opinion sentiments in a user’s tweets. The opinion contradiction level is defined as follows:

$$C_u = \frac{\nu \cdot \sigma_u^2}{\nu + (\mu_u)^2} W_u,$$

where $\nu$ is a normalizing constant, and $W_u$ is a weight function to compensate for the contradiction value of users’ varying number of tweets. In

$$W_u = \left(1 + \exp\left(\frac{|S^{(u)}| - \bar{S}^{(u)}}{\beta}\right)\right)^{-1},$$

$|S^{(u)}|$ is the number of $u$’s tweets, $|S^{(u)}|$ is the average number of all users’ tweets, and $\beta$ is a scaling factor. We set $\nu = 1$ and $\beta = 6$ experimentally.

Figure 3 shows a composite of user volume (blue bars) and prediction accuracy (colored lines). Users are grouped according to the score of sentiment contradiction ($x$-axis). For example, the bar between 0 and 0.05 represents the percentage of users with a contradiction level between
0 and 0.05. For each user group, the prediction accuracies of all methods on this group are displayed.

Over the four topics, TIM beats other methods on all user groups. Especially on user groups with higher levels of sentiment contradiction (that is, the last two groups for each topic), all unified methods, that is, TIM, TIM_S1, TIM_S2, and SVM, have significantly better performance compared to individual-based methods. This clearly shows the effectiveness of social influence on those users who are easily influenced. Among the three unified methods, TIM performs the best. It verifies that our proposed method could learn interpersonal influence more accurately.

For users at the lowest contradiction level (the first group from the left), individual-based methods and unified models all achieve about 85 percent accuracy, indicating that these users are difficult to influence, and their behaviors are relatively stable and thus predictable. Knowing this, we might not want to spend much on these stubborn users in real marketing activities. Instead, we should focus on how to influence those customers with high sentiment contradiction. For each topic in our dataset, about 15 percent of users fall into this category, and the proposed TIM method has the best prediction ability for this category of users.

In the future, we’ll consider hidden Markov models to comply with various other social scenarios, such as song recommendations on last.fm.

Acknowledgments
This work was submitted on 1 December 2015, first revised on 3 May 2016, and finally revised on 23 August 2016. It was supported by the Research Grants Council of Hong Kong (PolyU S502/12E, PolyU 152094/14E), National Natural Science Foundation of China (61272291), and Hong Kong Polytechnic University (4-BCB5, G-YBJP).

References