Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence

Will You "Reconsume" the Near Past? Fast Prediction on Short-Term Reconsumption Behaviors *

Jun Chen, Chaokun Wang and **Jianmin Wang**
School of Software, Tsinghua University, Beijing 100084, P.R. China.
chenjun14@mails.thu.edu.cn, {chaokun, jimwang}@tsinghua.edu.cn

Abstract

The short-term reconsumption behaviors, i.e. "reconsume" the near past, account for a large proportion of people's activities every day and everywhere. In this paper, we firstly derived four generic features which influence people's short-term reconsumption behaviors. These features were extracted with respect to different roles in the process of reconsumption behaviors, i.e. users, items and interactions. Then, we brought forward two fast algorithms with the linear and the quadratic kernels to predict whether a user will perform a short-term reconsumption at a specific time given the context. The experimental results show that our proposed algorithms are more accurate in the prediction tasks compared with the baselines. Meanwhile, the time complexity of online prediction of our algorithms is $O(1)$, which enables fast prediction in real-world scenarios. The prediction contributes to more intelligent decision-making, e.g. potential revisited customer identification, personalized recommendation, and information re-finding.

Introduction

People's reconsumption behaviors exist everywhere and happen every day. We loop our favourite songs while searching for the newly released music tracks from our favourite artists. We eat regularly at our favourite restaurants while experiencing fresh tastes at the new restaurants recommended by our friends. People's consumptions of items, e.g., songs, restaurants and web pages, consist of both novelty-seeking behaviors and reconsumption behaviors which are alternated now and then (Anderson et al. 2014). Thus, an interesting question is: Are these seemingly random behaviors, especially the reconsumption behaviors predictable?

(Anderson et al. 2014) proposed a method to predict what will most likely be reconsumed given the premise that the user is surely about to perform a reconsumption. However, this assumption leaves the more fundamental problem blank, i.e. whether or not a user will reconsume at a specific time. Answering the "whether" question is no inferior than answering the "what" question. If whether a user will reconsume could be predicted, then different data processing strategies would be used by systems, e.g., web revisitation or novel web exploration, recommendation from the past or from the unobserved. The "whether" problem can be considered as a *switch* that opens the doors of two disjoint problems and narrows the problem domains therein.

The prediction on whether or not a user will reconsume at a specific time is of broad interest. In recommendation tasks (Bobadilla et al. 2013), the prediction results help recommender systems understand whether the novel unobserved items or the already consumed items are more appropriate to be recommended at a specific time. In web browsers, the prediction results suggest whether or not the visited web pages should be cached for later revisitation (Adar, Teevan, and Dumais 2008). In commercial business, the prediction results enable restaurants and supermarkets to identify potential revisited customers who are further delivered with coupons (Han, Back, and Barrett 2009). Furthermore, inside mobile phones, the prediction could also help the system decide whether or not the pre-launching of that application is necessary out of the considerations for efficiency and user friendliness (Xia and Lam 2012).

Similar to the taxonomy of web revisitations by Adar *et al*. (Adar, Teevan, and Dumais 2008), the general reconsumption behaviors of human can also be classified into the short-term reconsumptions (fast group), the medium-term reconsumptions (medium group) and the long-term reconsumptions (slow group). In this paper, we mainly focus on the study of the **S**hort-**T**erm **REC**onsumption behaviors, *abbr*. STREC behaviors (i.e., reconsume the near past). The dynamics of STREC behaviors make it more difficult and important to study compared with the long-term and the medium-term reconsumption behaviors. The prediction of people's STREC behaviors is nontrivial and is confronted with a number of challenges: First, STREC behaviors are more prone to change with time compared with long-term reconsumption behaviors, which makes it more difficult to model the dynamics. Second, STREC behaviors are seemingly random and they are likely to be influenced by multiple factors like users, items and user-item interactions. Third, people may have different STREC behavior preferences in different domains, and it is difficult to extract general domain-independent patterns to represent people's STREC behaviors.

In this paper, we studied people's STREC behaviors in a

*Corresponding author: Chaokun Wang.

domain-independent way. Four prominent features were extracted given the user's recent consumption history. Two fast algorithms were proposed by solving an optimization problem to predict whether or not a user will perform a reconsumption at a specific time based on these features.

We summarize our main contributions as follows:

- We analyzed the major domain-independent factors in influencing people's STREC behaviors. The factors in the analysis covered the user aspect, the item aspect and the interactional aspect between users and items.
- The problem of the binary prediction on people's STREC behaviors, i.e., whether or not a user will perform a reconsumption given the user's consumption history, was presented and formulated. To the best of our knowledge, the problem has not been studied in a domain-independent way in the literature.
- Two fast prediction algorithms were proposed based on these domain-independent features to address the above problem. The proposed methods can be further improved by simply appending more domain-specific features to our feature vectors.
- The experimental results showed that the proposed methods demonstrated promising effectiveness and efficiency in the prediction. They outperformed the reference approaches using the state-of-the-art classifiers in the prediction experiment.

Related Work

People's reconsumption behaviors are observed frequently in many fields. We introduced several studies on people's reconsumption behaviors in this section.

Web Revisitation. As a major kind of reconsumption types, the web revisitation has been well studied (Adar, Teevan, and Dumais 2008; Liu et al. 2012; Yu et al. 2013), where people revisit the web pages that have already been visited before. People usually conduct web revisitation through the browsers' URL auto-completion, bookmarks, history sidebars, back and forward buttons (Kawase, Papadakis, and Herder 2011; Kawase, Herder, and Nejdl 2011). Zhang and Zhao (Zhang and Zhao 2011) found that about 39.3% of web page views are revisitations, and the proportion could be higher (about 59.6%) when considering the tab-switches in web browsers as revisitations. Generally, people tend to revisit the popular web pages and those which have recently been viewed (Catledge and Pitkow 1995).

Repeat Queries. Repeat queries, a.k.a. reconsumption of web queries, account for a large amount of web search traffic (Tyler and Teevan 2010). For example, about 40% web queries are repeat queries in the Yahoo's searching logs (Teevan et al. 2007). People frequently type the same queries in the web search engine to re-find exactly the same information they have viewed before or to follow the updated information on the topics they have explored (Teevan et al. 2007). Compared with the repeat queries in web search, the query repetition rate is reported higher (55.76%) in the short-text corpus — Twitter (Teevan, Ramage, and Morris 2011). People are likely to use repeat queries on Twitter to monitor topics over time.

Information Re-finding. Information re-finding is another challenging topic in people's reconsumption behaviors, where people attempt to re-find the information they have come across before. One of the most common scenarios of information re-finding is to re-find past emails (Elsweiler, Baillie, and Ruthven 2011). It is reported that about 55% selections on messages in email clients are re-finding behaviors (Elsweiler, Harvey, and Hacker 2011). (Elsweiler, Harvey, and Hacker 2011) used a generalized linear model to identify if the sequence of operations in the email client contains a re-finding behavior. The problem they have addressed is different from ours in this paper. They focused on identifying re-finding behaviors among the past consumptions rather than predicting the future, and they used the features related to operations in email clients, e.g., the selections on email entries and the opening of folders, instead of the domain-independent features used in this paper.

People's reconsumption behaviors are also observed in other fields like the repeat purchase (Chiu et al. 2012; Weisberg, Te'eni, and Arman 2011), the revisitation on restaurants from customers (Han, Back, and Barrett 2009; Park and Jang 2014) as well as the repeat social choice (Sarma et al. 2012). To the best of our knowledge, the most related work with ours is the analysis on the dynamics of reconsumptions (Anderson et al. 2014), in which the authors found that the quality of items and the recency effect have significant impacts on people's choice of items for reconsumption given the premise that the system already knows the user will perform a reconsumption. By contrast, our work attempts to predict whether a user will reconsume at a specific time, which is exactly the study on their premise. As far as we know, this is the first work addressing the domain-independent prediction on people's short-term reconsumption behaviors.

Preliminaries

In this work, we use a sliding window along the sequence of each user's consumptions ordered by time. The sliding window always maintains the k most recent consumptions for each user, and the length of the window k is fixed. As the user performs a new consumption, the sliding window takes a step forward. Thus, the sliding window also represents the updated user "context" to some extent. There are several basic concepts in our work:

Definition 1 *A consumption transaction t is a random variable representing an arbitrary item. The consumption history of user u is represented by a sequence of consumption transactions, $T_u = \{t_1^u, t_2^u, ..., t_{|T_u|}^u\}$.*

Definition 2 *A k-length sliding window, denoted as W_k, is a queue which maintains the k most recent consumption transactions of a user till now. Once the user issues a new transaction t on an arbitrary item, t is pushed into the tail of W_k. Meanwhile, if the number of transactions in W_k exceeds the limit k, the head transaction of W_k is popped.*

The sliding window W_k consists of no more than k consumption transactions. Based on the use of sliding window, we can formally define the STREC behaviors:

Definition 3 *Given the k-length sliding window W_k of a user by now and a new consumption transaction t, we call t a STREC behavior,* iff $t \in W_k$. *Otherwise, t is considered as a novel consumption.*

According to Definition 3, whether or not consumption transaction t is a STREC behavior is influenced by the setting of k. To investigate people's STREC behaviors, k is usually not large so that only the repetitions on the recent consumptions are considered. Thus, we formulate the problem of the binary prediction of STREC behaviors as below:

Definition 4 *Given the k-length sliding window $W_k^{u,t}$ of user u right before performing consumption transaction t, the problem of the* binary prediction of STREC behaviors *is to predict whether or not $t \in W_k^{u,t}$ where t is unknown.*

In this work, we use k-last-visited items to define sliding window rather than the elapsed time because the continuity makes the time bound hard to determine in real-world scenarios, and the k-last-visited items are representative enough for people's (re)consumption contexts.

Binary STREC Behavior Prediction

In this section, we first introduce the four generic features that we propose to represent user's reconsumption context. Then, two fast prediction methods are discussed.

Feature Extraction

So as to model people's willingness to reconsume, we propose the following features: (1) item popularity, (2) item reconsumption ratio, (3) user reconsumption ratio, and (4) window repeat ratio. These features correspond to the three major aspects regarding STREC behaviors, i.e. item aspect (1-2), user aspect (3) and interactional aspect (4).

Item Popularity The quality of items has an important effect upon the reconsumption behaviors (Anderson et al. 2014), and the items with high quality are usually much more likely to be reconsumed. Item popularity (*abbr.* IP) is an ideal measurement of the quality of each item.

Let $freq(x)$ be the frequency of item x in the given data set. We measure the popularity of item x as its fraction of the maximum item frequency in the data set,

$$h_{IP}(x) = \frac{log(1 + freq(x))}{\max_{y \in \mathbf{X}} log(1 + freq(y))}, \quad (1)$$

where $\mathbf{X}$ is the item set. The log operator is used to adjust the skewed distribution of item popularity. To represent the given sliding window W_k, we used the average popularity:

$$h_{IP}(W_k) = \frac{1}{|W_k|} \sum_{x \in W_k} h_{IP}(x). \quad (2)$$

By analyzing the relationship between $h_{IP}(W_k)$ and people's reconsumption willingness, we find that the probability of reconsumption increases as the average item popularity of sliding window increases.

Item Reconsumption Ratio As another way to explore the intrinsic item factors upon STREC behaviors, it is a straightforward intuition that the different reconsumption ratios of items also lead to different probabilities to perform reconsumptions. The absolute item reconsumption ratio is defined as its probability to be observed as a reconsumption along the user consumption transaction sequences,

$$h_{AIRR}(x) = log(1 + \frac{\sum_{u \in \mathbf{U}} \sum_{t \in T_u} \mathbb{1}_{t=x \wedge t \in C_k^{u,t}}}{\sum_{u \in \mathbf{U}} \sum_{t \in T_u} \mathbb{1}_{t=x}}), \quad (3)$$

where $\mathbf{U}$ is the user set, T_u is the consumption transaction sequence of user u, $\mathbb{1}_{cond}$ is the indicator function which returns 1 when $cond$ is satisfied, and otherwise returns 0. Similar to the definition of the item popularity, the value of the absolute item reconsumption ratio is also adjusted by a log operator.

The relative item reconsumption ratio is defined as the fraction of the maximum absolute item reconsumption ratio in the data set,

$$h_{IRR}(x) = \frac{h_{AIRR}(x)}{\max_{y \in \mathbf{X}} h_{AIRR}(y)}. \quad (4)$$

We use the average item reconsumption ratio (*abbr.* IRR) of a sliding window as another metric,

$$h_{IRR}(W_k) = \frac{1}{|W_k|} \sum_{x \in W_k} h_{IRR}(x). \quad (5)$$

In our experiments, we found that people's reconsumption probability is also increasing as the average item reconsumption ratio of sliding window gets larger.

User Reconsumption Ratio The user preference plays a key role in people's behaviors (Zhang, Wang, and Wang 2014), and the preferences are usually diverse among the crowds (Jung, Hong, and Kim 2005). The user factor is absolutely an important signal in analyzing people's STREC behaviors. For example, many people have brand loyalty and re-purchase on a few brands frequently (Chiu et al. 2012).

However, it is nontrivial to model every user aspect related to the personalities and preferences. We use the simple and straightforward feature of the user factor, i.e., user reconsumption ratio, to analyze how STREC behaviors can be affected by different user personalities and preferences. The ratio is defined as the probability that a user performs a reconsumption along her consumption transaction sequence,

$$h_{URR}(u) = \frac{\sum_{t \in T_u} \mathbb{1}_{t \in W_k^{u,t}}}{|T_u|}. \quad (6)$$

User reconsumption ratio (*abbr.* URR) is not time-aware, and it should be a static and intrinsic feature of user. A larger value of this feature indicates that the user is more likely to perform a reconsumption.

Window Repeat Ratio Besides the item factors and the user factor, we also explored the impact of the interactional factor upon people's STREC behaviors. Given a k-length sliding window W_k, we would like to know if the total number of reconsumption times in W_k influences people's reconsumption behavior in the next step. This feature is based on the hypotheses that people may get satiated with reconsuming several items again and again, or on the contrary, the compact reconsumptions increase the probability to perform another reconsumption again like the "Mathew Effect" (Merton 1968) in people's STREC behaviors.

Let $DS(W_k)$ be the set of *distinct* items in W_k, and $1 \leq |DS(W_k)| \leq k$. We use the proportion of the reconsumptions in the current sliding window to measure the window repeat ratio (*abbr.* WRR),

$$h_{WRR}(W_k) = 1 - \frac{|DS(W_k)|}{k}. \tag{7}$$

Obviously, $0 \leq h_{WRR}(W_k) < 1$. The larger the value of $h_{WRR}(W_k)$ is, the more compact the reconsumptions are in the sliding window. In our experiments, we found that the reconsumption probability is almost linear to the value of window repeat ratio in the current sliding window. This observation meets the hypothesis that "the rich get richer" effect also exists in people's STREC behaviors.

Fast Prediction Methods

In this paper, we attempt to combine all the four generic features to predict people's STREC behaviors. Given the k-length sliding window $W_k^{u,t}$ of user u right before performing consumption transaction t, we construct a vector $\mathbf{x}_{u,t} = \{h_{IP}(W_k^{u,t}), h_{IRR}(W_k^{u,t}), h_{URR}(u), h_{WRR}(W_k^{u,t})\}^T$ combining all the factors mentioned above together. Next, we propose two fast binary prediction methods with the linear and the quadratic kernels, respectively, to address our problem in this paper.

Linear Method Firstly, we use a linear hyperplane to separate the points in the 4-dimensional feature space. Our linear method predicts the probability of u's new consumption transaction t to be a reconsumption with the probability $\Pr_{\mathcal{L}}(u,t) = \mathbf{w}^T\mathbf{x}_{u,t},\ s.t., \sum_i \mathbf{w}_i = 1$. $\mathbf{x}_{u,t}$ is the aforementioned feature vector which represents the user's consumption "context", and $\mathbf{w}$ is the vector representing the hyperplane. So as to get the optimal hyperplane, it naturally leads to the following optimization problem,

$$\begin{aligned} \underset{\mathbf{w}^*_{\mathcal{L}}}{\operatorname{argmin}}\ \mathcal{L}(\mathbf{w}) &= \sum_{u\in\mathbf{U}}\sum_{t\in T_u}(\mathbf{w}^T\mathbf{x}_{u,t} - \mathbb{1}_{t\in W_k^{u,t}})^2, \\ s.t. &\sum_i \mathbf{w}_i = 1, \end{aligned} \tag{8}$$

where $W_k^{u,t}$ is the k-length sliding window of user u just before performing a new unknown consumption transaction t. Since the value of each dimension in $\mathbf{x}_{u,t}$ is between 0 and 1, the value of $\mathbf{w}^T\mathbf{x}_{u,t}$ will therefore be limited in the range $[0.0, 1.0]$. To solve this optimization problem, we rewrite the objective function $\tilde{\mathcal{L}}$ by adding a Lagrange multiplier,

$$\underset{\mathbf{w}^*_{\tilde{\mathcal{L}}}}{\operatorname{argmin}}\ \tilde{\mathcal{L}}(\mathbf{w}) = \sum_{u\in\mathbf{U}}\sum_{t\in T_u}(\mathbf{w}^T\mathbf{x}_{u,t} - \mathbb{1}_{t\in W_k^{u,t}})^2 + \lambda\sum_i \mathbf{w}_i. \tag{9}$$

We solve this problem using the gradient method by updating the hyperplane vector step by step using the following updating rule,

$$\frac{\partial\tilde{\mathcal{L}}(\mathbf{w})}{\partial\mathbf{w}_i} = \sum_{u\in\mathbf{U}}\sum_{t\in T_u} 2\mathbf{x}_{u,t,i}(\mathbf{w}^T\mathbf{x}_{u,t} - \mathbb{1}_{t\in W_k^{u,t}}) + \lambda. \tag{10}$$

Our linear method works by separating the points representing reconsumptions and novel consumptions in the 4-dimensional feature space with a linear hyperplane. If $\Pr_{\mathcal{L}}(u,t) > \frac{1}{2}$ for user u just before performing a new consumption transaction t, our linear method will predict t to be a reconsumption.

Quadratic Method In our experiments of the four generic features, we observed that people would be more likely to perform a reconsumption if any one of the four features is large enough. Thus, we believe it is better to separate the feature points using a hypersphere like a four dimensional ellipsoid. The probability of user u's new consumption transaction t being a reconsumption should be $\Pr_{\mathcal{Q}}(u,t) = \sqrt{\mathbf{w}^T diag(\mathbf{x}_{u,t})^2\mathbf{w}}$. We bring forward our quadratic method by solving the following quadratic optimization problem:

$$\begin{aligned} \underset{\mathbf{w}^*_{\mathcal{Q}}}{\operatorname{argmin}}\ \mathcal{Q}(\mathbf{w}) &= \sum_{u\in\mathbf{U}}\sum_{t\in T_u}(\sqrt{\mathbf{w}^T diag(\mathbf{x}_{u,t})^2\mathbf{w}} - \mathbb{1}_{t\in W_k^{u,t}})^2, \\ & s.t. \mathbf{w}^T\mathbf{w} = 1, \end{aligned} \tag{11}$$

where $diag(\mathbf{x}_{u,t})$ is the diagonal matrix representation of vector $\mathbf{x}_{u,t}$. Similarly, this optimization problem can be solved by adding the Lagrange multiplier to the new objective function $\tilde{\mathcal{Q}}(\mathbf{w})$,

$$\begin{aligned} \underset{\mathbf{w}^*_{\tilde{\mathcal{Q}}}}{\operatorname{argmin}}\ \tilde{\mathcal{Q}}(\mathbf{w}) &= \sum_{u\in\mathbf{U}}\sum_{t\in T_u}(\sqrt{\mathbf{w}^T diag(\mathbf{x}_{u,t})^2\mathbf{w}} - \mathbb{1}_{t\in W_k^{u,t}})^2 \\ &+ \lambda\mathbf{w}^T\mathbf{w}. \end{aligned} \tag{12}$$

The vector hypersphere is updated using the following rule,

$$\begin{aligned} \frac{\partial\tilde{\mathcal{Q}}(\mathbf{w})}{\partial\mathbf{w}_i} &= \sum_{u\in\mathbf{U}}\sum_{t\in T_u} 2(\sqrt{\mathbf{w}^T diag(\mathbf{x}_{u,t})^2\mathbf{w}} - \mathbb{1}_{t\in W_k^{u,t}}) \\ &\frac{2\mathbf{x}^2_{u,t,i}\mathbf{w}_i}{\sqrt{\mathbf{w}^T diag(\mathbf{x}_{u,t})^2\mathbf{w}}} + 2\lambda\mathbf{w}_i \\ &= \sum_{u\in\mathbf{U}}\sum_{t\in T_u} 4\mathbf{x}^2_{u,t,i}\mathbf{w}_i(1 - \frac{\mathbb{1}_{t\in W_k^{u,t}}}{\sqrt{\mathbf{w}^T diag(\mathbf{x}_{u,t})^2\mathbf{w}}}) \\ &+ 2\lambda\mathbf{w}_i. \end{aligned} \tag{13}$$

If $\Pr_{\mathcal{Q}}(u,t) > \frac{1}{2}$ for user u just before performing a new consumption transaction t, our quadratic method will predict t to be a reconsumption.

The proposed methods are efficient in the tasks of the binary prediction of STREC behaviors. The time complexity of online computation of four features and the prediction probability is $O(1)$ since a priority queue is used as the sliding window moves forward and the values can be obtained

Table 1: Statistics of the data sets.

Data Set	#.Users	#.Items	#.Transactions
Lastfm	992	964,463	16,986,614
BrightKite	51,406	772,966	4,747,281
Gowalla	107,092	1,280,969	6,442,892
ManicTime	44	22,808	253,283

by simply subtracting old ones and adding new ones. There is no need to re-compute these features by enumerating the window again online. In contrast, the only time-consuming part is to learn the hyperplane and the hypersphere in our methods, which however, can be finished and updated offline in a batch mode periodically. Thus, our methods are very efficient in the binary prediction.

Experiments

Data Sets

We used four data sets to evaluate the performance of our prediction algorithms. The statistics of these data sets are shown in Table 1.

Lastfm. This publicly available data set is a collection of people's listening records on Last.fm (Celma 2010). People listen to music frequently in their daily lives. The listening records are the mixture of novel songs and loops of previously heard songs.

BrightKite. This publicly available data set contains people's check-ins at different locations (Cho, Myers, and Leskovec 2011). People's repeat check-ins at the same location which is recognized by the location-id or the coordinates, can be considered as reconsumptions.

Gowalla. Gowalla was a location-based social network available. This data set contains users' check-ins of locations (Cho, Myers, and Leskovec 2011). Similar to BrightKite, people's repeat check-ins at the same location are considered as reconsumptions. However, Gowalla has nearly twice the number of users and items compared with BrightKite, and this is also the sparsest data set in our experiments. Besides, the overall reconsumption ratio of Gowalla is very different from that of BrightKite (see Fig. 1).

ManicTime. To broaden the types of reconsumption data in our analysis, we also collected a new data set which contains the using logs of people's desktop applications on PCs. We recruited 44 college students to conduct the rewarding experiments. Each of the volunteers installed the time management tool — ManicTime [1], on their personal computers. This tool monitors the using logs like the launching time, the shutdown time and the lasting period of each application without interrupting the daily work of users. All the volunteers agreed to keep this tool running as a background service and collect logs for at least one month. They are assured that the users will be anonymized and the data will be used only for research purpose. In this data collection, the repeat use on the same application by a same user is considered as reconsumptions.

[1] http://www.manictime.com

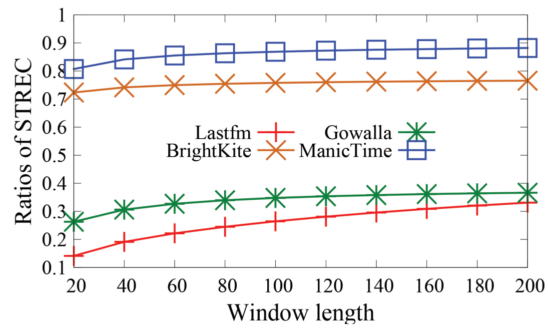


Figure 1: The ratios of the STREC behaviors of four data sets under different sliding window length.

Experimental Settings

To the best of our knowledge, there is no prior work directly addressing the problem of predicting whether a user will perform a reconsumption. Therefore, we used the state-of-the-art classifiers — the support vector machine (SVM) and the discriminant analysis (DA) using our proposed features as the comparisons to our proposed methods in this paper. Since we presented both the linear and the quadratic methods, we also evaluated the SVM and DA with the linear and the quadratic kernel functions, and they are denoted as SVM-L, SVM-Q, DA-L and DA-Q, respectively. The parameters in SVM and DA are tuned in advance to maximize their prediction accuracy on each data set. Besides, our proposed methods are denoted by ST-L (linear method) and ST-Q (quadratic method), respectively.

STREC Behavior Prediction

According to Definition 3, whether or not a consumption transaction is considered as a STREC behavior can be affected by the length of sliding window. Windows of different lengths lead to different definitions of STREC behaviors as well as different ground truth in our evaluations on the same data sets. Thus, we firstly compared these prediction methods under the same setting of window length.

Table 2 shows the results of predicting whether or not a new transaction is a STREC behavior in all data sets when the window length $k = 20$. TP, TN, FN and FP represent True-Positive, True-Negative, False-Negative and False-Positive, respectively. Thus, the values of $\frac{\text{TP}}{\text{TP+TN}}$ and $\frac{\text{FN}}{\text{FN+FP}}$ can be used to measure the ability of the methods to identify the reconsumption behaviors and the novel consumption behaviors, respectively. Furthermore, the overall prediction accuracy of these methods can be evaluated by $\frac{\text{TP+FN}}{\text{TP+TN+FN+FP}}$.

We can see that ST-L and ST-Q have the dominating overall prediction accuracy on all the four data sets and their performance is stable and promising. ST-Q has the highest prediction accuracy 0.7741 in the BrightKite set, and ST-L outperforms the other methods in the Lastfm, the Gowalla and the ManicTime sets by reaching the prediction accuracy, 0.8799, 0.7526 and 0.8090, respectively. The limitation of our proposed methods is the hypothesis that the fea-

Table 2: Results of predicting reconsumption behaviors. TP, TN, FN, FP represent True-Positive, True-Negative, False-Negative, False-Positive, respectively. (k=20.)

(a) The Lastfm set.

Method	TP/(TP+TN)	FN/(FN+FP)	(TP+FN)/(TP+TN+FN+FP)
DA-L	0.6095	0.9098	0.8662
DA-Q	**0.6321**	0.8762	0.8407
SVM-L	0.0	**1.0**	0.8548
SVM-Q	0.0	**1.0**	0.8548
ST-L	0.3467	0.9704	**0.8799**
ST-Q	0.3078	0.9762	0.8791

(b) The BrightKite set.

Method	TP/(TP+TN)	FN/(FN+FP)	(TP+FN)/(TP+TN+FN+FP)
DA-L	0.6237	0.8650	0.6985
DA-Q	0.5773	**0.8927**	0.6751
SVM-L	0.7409	0.7110	0.7316
SVM-Q	0.9181	0.1772	0.6885
ST-L	0.8840	0.5282	0.7738
ST-Q	0.8649	0.5719	**0.7741**

(c) The Gowalla set.

Method	TP/(TP+TN)	FN/(FN+FP)	(TP+FN)/(TP+TN+FN+FP)
DA-L	0.5901	0.7523	0.7084
DA-Q	0.6029	0.7246	0.6916
SVM-L	0.4351	0.8444	0.7336
SVM-Q	**0.9968**	0.0008	0.2704
ST-L	0.2978	0.9214	**0.7526**
ST-Q	0.2975	0.9210	0.7523

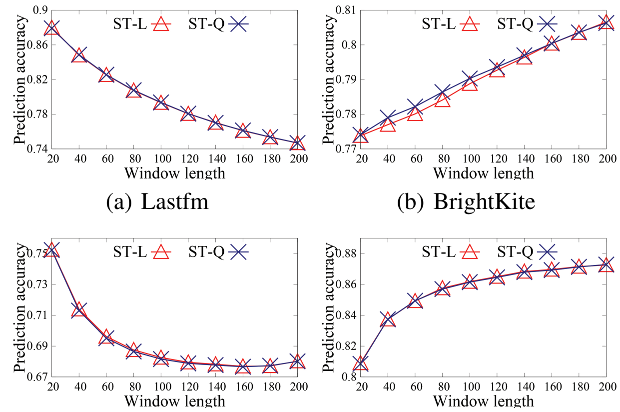
(d) The ManicTime set.

Method	TP/(TP+TN)	FN/(FN+FP)	(TP+FN)/(TP+TN+FN+FP)
DA-L	0.7585	0.5485	0.7157
DA-Q	0.7824	0.5082	0.7265
SVM-L	0.0	**1.0**	0.2038
SVM-Q	**0.9949**	0.0531	0.8030
ST-L	0.9731	0.1681	**0.8090**
ST-Q	0.9767	0.1514	0.8085

ture points in the proposed 4-dimensional space are separable. According to the results of the prediction experiments, both of the separations worked effectively. Meanwhile, we also observed that the SVM methods are prone to over-learn and lead to biased prediction results, e.g. either all positive or all negative. By contrast, our methods could balance the prediction results better.

Impacts of Window Length

People's consumption transactions may or may not be identified as STREC behaviors under different settings of window length according to Definition 3. Fig. 1 illustrates the overall ratios of STREC behaviors with respect to different settings of window length of the four data sets. The changes of STREC behavior ratio of BrightKite, Gowalla and ManicTime, are all very smooth as the window length increases, which means most people are not likely to reconsume what have been consumed far from now. This observation is also a proof that the recency effect is significant



(a) Lastfm (b) BrightKite

(c) Gowalla (d) ManicTime

Figure 2: The prediction accuracy under different settings of window length of the four data sets.

in the location check-ins and the use of desktop pplications. In contrast, the STREC ratio of Lastfm increases quickly from 0.14 ($k = 20$) to 0.33 ($k = 200$). It means we are likely to reconsume a music track that was heard long ago, which may result from that we usually have a playlist when listening to music, and only perform reconsumptions after listening through the whole playlist.

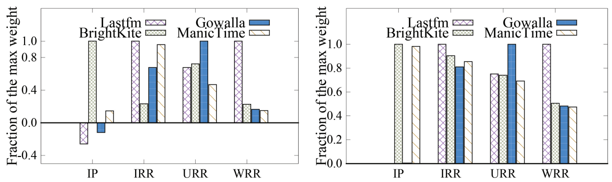
The accuracy of binary STREC predictions under different settings of window length is illustrated in Fig. 2. We can see that the accuracy performance of ST-L and ST-Q is very similar regardless of the change of window length. In addition, the accuracy performance of our methods is also subject to the overall reconsumption ratios as illustrated in Fig. 1. It will be the most difficult (with the least accuracy) to predict STREC behaviors if the overall reconsumption ratio is around 50% under a certain setting of window length, and vice versa.

Analysis of Learned Parameters

Next, we analyzed the parameters of our methods obtained by solving the optimization problems to study the relative importance of each generic feature. The parameters of the hyperplane and the hypersphere are shown in Fig. 3. The values of parameters in each vector **w** are their fractions of the maximum value in **w**, i.e., $\frac{\mathbf{w}_i}{\max_j \mathbf{w}_j}$.

For the parameters of the linear method, we can see that the IRR and the WRR dimensions are significant in the binary prediction in the Lastfm set, while the IRR and the URR dimensions are both significant in the Gowalla and the ManicTime sets. In the BrightKite set, the most significant dimension is the IP by contrast. Different from the quadratic method, parameters of the linear method have the possibility to be negative, e.g., the IP dimension in Fig. 3(a).

For the parameters of the quadratic method, we observed similar trend in the BrightKite and the ManicTime that the importance of the IP, the IRR, the URR and the WRR dimensions is in descending order. Besides, the most significant dimension of the Gowalla set is the URR dimension,



(a) ST-L Parameters (b) ST-Q Parameters

Figure 3: The learned parameters of the hyperplane and the hypersphere of our methods. Parameters are shown as their fractions of the maximum value in each vector **w**. ($k = 20$)

while the IRR and the WRR dimensions seem to be the most important in the Lastfm set.

Based on the discussion on the experiments, we can see that our methods are effective and efficient in solving the binary prediction of STREC behaviors.

Conclusion and Future Work

In this paper, we addressed the problem of predicting whether or not people would perform a reconsumption at a specific time. Four generic features were derived based on people's recent (re)consumption behaviors. We also proposed two fast binary prediction algorithms w.r.t. the linear and the quadratic kernels. The experimental results showed that our methods are effective in the prediction compared with the reference methods using the state-of-the-art classifiers. In the future, we will expand our current work to further predict which item will most probably be reconsumed when our methods foresee the happening of a reconsumption behavior. Then, both the "*whether*" and "*what*" questions about people's reconsumption behaviors will be answered.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (No. 61373023, No. 61170064, No. 61133002) and the National High Technology Research and Development Program of China (No. 2013AA013204).

References

Adar, E.; Teevan, J.; and Dumais, S. T. 2008. Large scale analysis of web revisitation patterns. In *CHI*, 1197–1206.

Anderson, A.; Kumar, R.; Tomkins, A.; and Vassilvitskii, S. 2014. The dynamics of repeat consumption. In *WWW*, 419–430.

Bobadilla, J.; Ortega, F.; Hernando, A.; and Gutiérrez, A. 2013. Recommender systems survey. *Knowledge-Based Systems* 46:109–132.

Catledge, L. D., and Pitkow, J. E. 1995. Characterizing browsing strategies in the world-wide web. *Computer Networks and ISDN Systems* 27(6):1065–1073.

Celma, O. 2010. *Music Recommendation and Discovery in the Long Tail*. Springer. chapter 3.

Chiu, C.-M.; Hsu, M.-H.; Lai, H.; and Chang, C.-M. 2012. Re-examining the influence of trust on online repeat purchase intention: The moderating role of habit and its antecedents. *Decision Support Systems* 53:835–845.

Cho, E.; Myers, S. A.; and Leskovec, J. 2011. Friendship and mobility: User movement in location-based social networks. In *KDD*, 1082–1090.

Elsweiler, D.; Baillie, M.; and Ruthven, I. 2011. What makes re-finding information difficult? a study of email re-finding. In *ECIR*, 568–579.

Elsweiler, D.; Harvey, M.; and Hacker, M. 2011. Understanding re-finding behavior in naturalistic email interaction logs. In *SIGIR*, 35–44.

Han, H.; Back, K.-J.; and Barrett, B. 2009. Influencing factors on restaurant customers' revisit intention: The roles of emotions and switching barriers. *International Journal of Hospitality Management* 28(4):563–572.

Jung, S. Y.; Hong, J.-H.; and Kim, T.-S. 2005. A statistical model for user preference. *TKDE* 17(6):834–843.

Kawase, R.; Herder, E.; and Nejdl, W. 2011. Beyond the usual suspects: context-aware revisitation support. In *ACM conference on Hypertext and hypermedia*, 27–36.

Kawase, R.; Papadakis, G.; and Herder, E. 2011. Supporting revisitation with contextual suggestions. In *JCDL*, 227–230.

Liu, J.; Yu, C.; Xu, W.; and Shi, Y. 2012. Clustering web pages to facilitate revisitation on mobile devices. In *IUI*, 249–252.

Merton, R. K. 1968. The mathew effect in science. *Science* 159(3810):56–63.

Park, J.-Y., and Jang, S. 2014. Revisit and satiation patterns: Are your restaurant customers satiated? *International Journal of Hospitality Management* 38:20–29.

Sarma, A. D.; Gollapudi, S.; Panigraphy, R.; and Zhang, L. 2012. Understanding cyclic trends in social choices. In *WSDM*, 593–602.

Teevan, J.; Adar, E.; Jones, R.; and Potts, M. 2007. Information re-retrieval: repeat queries in yahoo's logs. In *SIGIR*, 151–158.

Teevan, J.; Ramage, D.; and Morris, M. R. 2011. Twittersearch: a comparison of microblog search and web search. In *WSDM*, 35–44.

Tyler, S. K., and Teevan, J. 2010. Large scale query log analysis of re-finding. In *WSDM*, 191–200.

Weisberg, J.; Te'eni, D.; and Arman, L. 2011. Past purchase and intention to purchase in e-commerce: The mediation of social presence and trust. *Internet Research* 21(1):82–96.

Xia, B., and Lam, I. 2012. Adaptive termination and pre-launching policy for improving application startup time. US Patent App. 13/161,703.

Yu, C.; Balakrishnan, R.; Hinckley, K.; Moscovich, T.; and Shi, Y. 2013. Implicit bookmarking: Improving support for revisitation in within-document reading tasks. *Int. J. Human-Computer Studies* 71:303–320.

Zhang, H., and Zhao, S. 2011. Measuring web page revisitation in tabbed browsing. In *CHI*, 1831–1834.

Zhang, J.; Wang, C.; and Wang, J. 2014. Learning temporal dynamics of behavior propagation in social networks. In *AAAI*, 229–236.