2015 IEEE International Conference on Data Mining

Mining Indecisiveness in Customer Behaviors

Qi Liu[1], Xianyu Zeng[1], Chuanren Liu[2], Hengshu Zhu[3], Enhong Chen[1,*], Hui Xiong[4], Xing Xie[5]
[1]School of Computer Science and Technology, University of Science and Technology of China
qiliuql@ustc.edu.cn, zengxy@mail.ustc.edu.cn, cheneh@ustc.edu.cn
[2] Decision Sciences & MIS Department, Drexel University, chuanren.liu@drexel.edu
[3] Big Data Lab, Baidu Research, zhuhengshu@baidu.com
[4]Rutgers Business School, Rutgers University, hxiong@rutgers.edu
[5]Microsoft Research, xing.xie@microsoft.com

***Abstract*—In the retail market, the consumers' indecisiveness refers to the inability to make quick and assertive decisions when they choose among competing product options. Indeed, indecisiveness has been investigated in a number of fields, such as economics and psychology. However, these studies are usually based on the subjective customer survey data with some manually defined questions. Instead, in this paper, we provide a focused study on automatically mining indecisiveness in massive customer behaviors in online stores. Specifically, we first give a general definition to measure the observed indecisiveness in each behavior session. From these observed indecisiveness, we can learn the latent factors/reasons by a probabilistic factor-based model. These two factors are the indecisive indexes of the customers and the product bundles, respectively. Next, we demonstrate that this indecisiveness mining process could be useful in several potential applications, such as the competitive product detection and personalized product bundles recommendation. Finally, we perform extensive experiments on a large-scale behavioral logs of online customers in a distributed environment. The results reveal that our measurement of indecisiveness agrees with the common sense assessment, and the discoveries are useful in predicting customer behaviors and providing better recommendation services for both customers and online retailers.**

I. Introduction

Advances in retail information systems have enabled us to collect a massive amount of customer consumption records. These information-rich behavioral logs open the opportunities for mining personalized customer profiles and customer preferences [1], [2], and thus enabling the development of better customer services [3]–[5], e.g. personalized item assortments.

Along this line, there is a particular interest in understanding the decision making process of customers [6]–[8]. However, with the rapidly growing number of product choices, it becomes difficult for customers to make decisions. More and more people suffer from the problem of indecisiveness, which refers to the inability of making quick and assertive decisions when they choose among competing product options [8]. For instance, Figure 1 illustrates one difficult choice of Dr. John Krumm[1], when he was shopping for eyeglass frames. To make a right decision, he had to refer to the suggestions from his social friends on Facebook (Figure 1(a)). According to Dr. Krumm's conclusion, the second frame (choice "B") finally won most votes (Figure 1(b)). We can see that indecisiveness is actually a common problem. Meanwhile, each time of indecisiveness could help capture some characteristics (e.g. self-confidence and depression) of the customer and the cognitive states from this customer to the items [9]. Therefore, it is an important research topic for understanding an individual's indecisiveness in marketing, management and psychology. Indeed, researchers have conducted some studies on the related issues, such as the measurement of indecisiveness [9]–[12], the reasons for indecisiveness [7], [13]–[16] and the way to reduce the difficulty of decision-making [3], [6], [17]–[20].

(a) Dr. Krumm's query (b) The voting result

Fig. 1. An example of Indecisiveness from Dr. John Krumm.

In spite of the importance of the previous studies, the scientific literature on the understanding of indecisiveness is still limited, or even a clear definition of indecisiveness is lacking [10]. Usually, for measuring indecisiveness, some questions (e.g. "I make decisions quickly") have to be manually prepared by the psychologists or sociologists. Then, the exact answers of the individuals to these questions are used for judging their scales of indecisiveness [11], [12]. This questionnaire-based perspective of measurement is not only very subjective but also labor intensive, and the further conclusions will be biased and even misleading. It is better to automatically mine indecisiveness from customer's everyday behaviors without manual interventions, i.e. in a complete data-driven way. Towards this goal, there are several challenges or questions. Specifically, how to effectively define the scale of observed indecisiveness without any unified standards (questions)? How to precisely figure out the latent and personalized indecisive indexes for each individual and item (item bundle)? How to exploit the mined indecisiveness for better services?

To address the challenges mentioned above, in this paper, we provide a focused study on mining indecisiveness from a data-driven perspective. Along this line, we first analyze customer behaviors in online consumptions and introduce a general definition to automatically measure the scale of the observed indecisiveness in each behavior session. Based on the existing understanding from other domains and the availability of data, this definition considers three major characteristics of one behavior session, e.g. the length of behaviors. Then, we introduce a method for learning the latent indecisive indexes

*Corresponding Author.
[1]http://research.microsoft.com/en-us/um/people/jckrumm/

1550-4786/15 $31.00 © 2015 IEEE
DOI 10.1109/ICDM.2015.78

for both customers and item bundles by a probabilistic factor-based model. In this way, the scale of indecisiveness in terms of each customer and each item bundle can also be quantified. Next, we demonstrate that the identified indecisiveness is useful in several potential applications. For instance, for each indecisive customer, we could recommend the item bundles from which she will choose (consume). Also, for each retailer, we could figure out the competitive items. Note that given any specific definition of indecisiveness, the proposed framework can be adopt for mining customer's indecisiveness, so it is a general framework. Finally, we evaluate our methods on a large-scale behavioral logs of Tmall (http://www.tmall.com/) customers in a distributed environment using MapReduce. The extensive experimental results suggest that our measurement of indecisiveness is effective, and the identified indecisiveness could be used to predict customer's final consumption behaviors and recommend competitive items. The main contributions of this paper are summarized as the following:

- To the best of our knowledge, this is the first comprehensive attempt for understanding of human indecisiveness in a data-driven way, based on the massive behavioral data from millions of online customers.
- We provide a general framework for automatically measuring observed indecisiveness in each behavior session and computing the latent indecisive indexes for both customers and item bundles. The computation process is domain-independent and without manual interventions.
- We present in-depth analysis for the possible applications of the indecisiveness measurement. Specifically, we propose to apply indecisiveness for finding the competitive items for online retailers and recommending the consumption choice of each customer.

II. Related Work

We review the research from judgement and decision making, psychology, management, and marketing literatures to assemble what we already know about indecisiveness.

The first research direction on the study of indecisiveness is the measurement of indecisiveness based on behavioral characteristics [9]–[12]. Actually, the scientific literature on indecisiveness is rather limited and even a clear definition of indecisiveness is lacking [10]. Fortunately, several scales for indecisiveness have been constructed and then used to investigate distinctive features of indecisiveness. These scales, including the Frost and Shows' Indecisiveness Scale (FIS) [11], and the Germeijs and De Boeck' Indecisiveness Scale (GIS) [12], usually consist of sets of questions that are manually given by psychologists or sociologists. For instance, the questions may be "While making a decision, I feel certain" and "It takes a long time to weigh the pros and cons before making a decision" [12], and each question is formulated as a statement for which the subjects had to indicate the extent of agreement (e.g. on a 7-point scale). Then, the properties (e.g. reliability) of these indecisiveness scales are validated by the empirically conducted psychometric testing [9].

In the second direction, researchers focus on investigating the reasons (e.g. the cognitive elements) of indecisiveness. For instance, in today's market democracies, people face an ever-increasing number of options, and it is necessary to figure out the relationship between the choice overload and indecisiveness. Interestingly, by the meta-analysis of many experiments, Scheibehenne et al. [13] found that adverse effects due to an increase in the number of choice options are not very robust. Similarly, Cho et al. showed that the ease or difficulty of making a decision is not inherent in the choices themselves but has much to do with the state-of-mind of the decision-maker [14], and Polman et al. also proposed that the choices vary according to regulatory focus, such that people who make choices for themselves are prevention-focused, whereas people who make choices for others are promotion-focused [7]. However, Sela et al. found that the size of the candidate item set do have impact on the justification and decision processes of customers (e.g. shifts choice from vices to virtues) [15]. Furthermore, both Chernev et al. [16] and Broniarczyk et al. [21] conducted detailed analysis of the effects of the size of the candidate item set and the effects of some other factors (e.g. decision goal) that may lead to indecisiveness. Rassin et al. revealed that both the personal profiles (e.g. gender differences) [22] and the "safety" of the decision [8] are the reasons for indecisiveness. Recently, based on the dual-system theory of judgment, Dhar and Gorlin proposed a framework for comprehensively understanding the preference construction processes that underlie decision effects (task and context effects) identified in the literature [23].

Given the reasons for indecisiveness, researchers in the third direction want to help reduce the decision difficulty of people [3], [6], [17]–[20]. As noted by Brooks, indecisiveness is an individual difference that is thought to be relatively stable across time and situation, and the notion of turning people from indecisive to decisive with training or intervention is unlikely to be successful; however, changing decision behavior is a more realistic endeavor [17]. Along this line, recommender system seems to be a useful tool for handling information overload [4], e.g. by changing the item assortment [3] or providing recommendation signs (e.g., "Best Seller", "Award Winner") [19]. If the recommended items are still too similar, the diversification based on their latent features could be used to reduce decision difficulty [18]. Meanwhile, it is suggested that sound and rational decision-making also depends on the prior emotional processing [6] and the goals (e.g. a hedonic goal or a utilitarian goal) [20]. Thus, it is not easy to change the decision behavior of decision-makers.

Indeed, besides indecisiveness, there are still several types of psychological traits expressed by customers, for instance, novelty seeking [24], interest expansion [25] and serendipitous discovery [26] all give the customers exciting and satisfied feelings by going through unfamiliar experiences. An incisive understanding of these personal behaviors is not only essential to many scientific disciplines, but also critical for business success [1], [27], e.g. steering customer behaviors [5] and providing more accurate recommendations [28]. Luckily, with more and more customer data being collected, it becomes much easier to conduct in-depth studies.

TABLE I

A TOY EXAMPLE OF THE CUSTOMER BEHAVIORIAL RECORDS.

UserId	ItemId	CategoryId	Action	Timestamp
U_1	a	C_1	Click	2014-07-08 20:05:20
U_1	b	C_1	Click	2014-07-08 20:06:40
...	...	...	...	...
U_1	a	C_1	Cart	2014-07-08 20:13:55
U_1	b	C_1	Collect	2014-07-08 20:14:20
U_1	b	C_1	Buy	2014-07-08 20:14:38
U_2	f	C_2	Click	2014-07-09 10:21:13
U_2	f	C_2	Buy	2014-07-09 10:21:20

TABLE II

EXAMPLES OF THE CUSTOMER BEHAVIOR SESSIONS.

SessionId	UserId	CategoryId	Item Sequence	Item Bundle
S_1	U_1	C_1	a,b,b,a,b,b,a,a,b,b	{a,b}
S_2	U_1	C_1	a,c,a,c,e,b,e,e,c,e	{a,b,c,e}
S_3	U_1	C_2	f,g,f	{f,g}
S_4	U_2	C_1	a,b,c,a	{a,b,c}
S_5	U_2	C_3	h,d	{h,d}

III. INDECISIVENESS MINING

In this section, we propose a framework to explore the indecisiveness embodied in each customer's behaviors. First, we clarify the general format of the behavioral records. Then, we explain our measurement of the observed indecisiveness. Next, we present the way to learn the latent indecisive indexes. Finally, we show the potential applications.

A. Preliminaries

We aim to automatically mine indecisiveness for the customers with online behaviors recorded. A toy example of these records is given in Table I, where the customers' (e.g. U_1)[2] actions (both the buying actions and the historical actions in decision-making, e.g., click, collect and cart) on each item (e.g. a) are included. By preprocessing based on segmentations (e.g. segment if no actions are recorded during a time interval, e.g. 1 hour) [29], we could split the records of one customer into different sessions. Table II lists some session examples, and the notation of item bundle in Table II is defined as:

Definition 1 ***[Item Bundle]*** *Given a customer and one of her behavior session, the corresponding item bundle(s) store the set of different items in this session. Thus, in a session there are usually multiple item bundles, e.g. for the session with item sequence {a,b,c,a}, item bundles are {a,b,c}, {a,b},{b,c} and etc.*

For simplicity, Table II only shows the biggest bundle in each session, and we use it to stand for all other bundles. Thus, the larger the item bundle, the more items are included in this session, i.e. viewed by the customer. If several items often appear in different sessions, these items will make up a frequent item bundle, and these frequent item bundles could be mined by some algorithms from frequent pattern mining, e.g. the Apriori-like ones [30]. It is intuitive that the items in one frequent item bundle may be in a relationship of either competitive or cooperative. Actually, the cooperative relationship has been exploited in the existing marketing solutions (e.g. in Amazon, the items that are frequently bought together will be combined for better selling)[3]. In the customers' behavior (e.g. click) sessions, it is more often that the items appear together may compete against each other, and the problem of how to quantitatively measure the coopetition relationships in these item bundles is still open. From Table II, we can see that U_1 has the choice difficulty for the items in category C_1, while U_2 may be a much quicker decision-maker. Without loss of generality, we only consider the items in the same category for each session, and meanwhile, we do not explicitly distinguish the different types (e.g. click and cart) of customer behaviors/actions. At last, Table III lists some of the notations.

TABLE III

SEVERAL IMPORTANT MATHEMATICAL NOTATIONS.

Notations	Description
$U = \{U_1, U_2, ..., U_M\}$	the set of customers
$I = \{a, b, ..., m, n\}$	the set of items
$S = \{S_1, S_2, ..., S_N\}$	the set of sessions
$B = \{B_1, B_2, ..., B_P\}$	the set of frequent item bundles
$D^S = \{D^S_{1,1}, D^S_{1,2}, ..., D^S_{M,N}\}$	the observed indecisiveness
$Y = \{Y_1, Y_2, ..., Y_M\}$	the latent indecisive index of each customer
$Z^S = \{Z^S_1, Z^S_2, ..., Z^S_N\}$	the latent indecisive index of each session
$Z^B = \{Z^B_1, Z^B_2, ..., Z^B_P\}$	the latent indecisive index of each item bundle
$S_B(j) = \{B_{j(p_1)}, B_{j(p_2)}, ..., B_{j(p_T)}\}$	the set of frequent item bundles existing in session S_j
$W_B(j) = \{w_{j(p_1)}, w_{j(p_2)}, ..., w_{j(p_T)}\}$	the weight of the frequent item bundles in session S_j
$Q = \{Q_1, Q_2, ..., Q_M\}$	the set of customer feature vectors
$V = \{V_1, V_2, ..., V_P\}$	the set of item bundle feature vectors

B. Observed Indecisiveness

In this subsection, we give a formal definition to measure the observed indecisiveness from each behavior session. Note that the measurement is domain-independent, which means it also works for other indecision, e.g. career indecision [12].

According to the measurement in psychology, indecisiveness could be represented by both longer decision and increased search for information [9], [11]. Based on existing understanding of indecisiveness and the availability of online customer data, we mainly focus on three characteristics and assumptions when measuring indecisiveness shown from one session: 1) The more actions in this session, the higher indecisiveness of the corresponding customer ("takes a long time" or "delaying"); 2) The more balanced distribution of the actions on different items, the higher indecisiveness ("feeling uncertain"); 3) The more transitions between the actions of different items, the higher indecisiveness ("reconsideration") [12]. Thus, we could define the observed indecisiveness in each behavior session as a function of three characteristics:

Definition 2 (Observed Indecisiveness) *Given a customer U_i and one of her behavior session S_j where the corresponding items form a bundle B_k, we define the observed indecisiveness of this session as*

$$D^S_{ij} = F(Length(S_j), Entropy(S_j), Trans(S_j)), \qquad (1)$$

where $F(.)$ can be any workable function, and $Length(S_j)$ denotes the $log(\cdot)$ of the number of actions, $Entropy(S_j)$ equals to $-\sum_{a \in B_k} p_j(a) log(p_j(a))/log(|B_k|)$ measures the distribution of the actions on different items (here, $p_j(a)$ is the probability that item a shows up in session S_j) and $Trans(S_j)$ is the number of transitions between different items divided by the total number of transitions. For instance, the $Length(S_j)$, $Entropy(S_j)$ and $Trans(S_j)$ of session S_1 in Table II are $log(10)$, $(-0.4*log(0.4)-0.6*log(0.6))/log(2)$ and $5/9$, respectively.

If let $F(Length(S_j), Entropy(S_j), Trans(S_j))$ denote a simple multiply of these elements, i.e. $Length(S_j) * Entropy(S_j) * Trans(S_j)$, then the observed indecisiveness of session S_1 is about 0.53[4]. In this way, we can measure the customer

[2] We use the expressions of "customer" and "user" interchangeably.

[3] http://en.wikipedia.org/wiki/Product_bundling

[4] Before computation, $Length(S_j)$ could be normalized into $[0, 1]$.

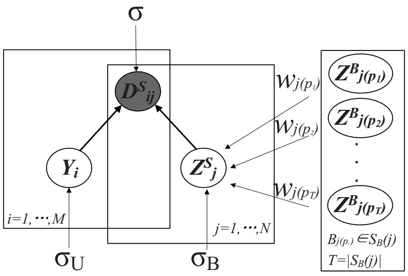


Fig. 2. The learning model for latent indecisive indexes.

indecisiveness for each scenario/domain as long as the decision making process are recorded. According to the above definition, the observed indecisiveness can locate in any range, e.g. in $(0, +\infty)$, and we could also normalize it into the range of $[0, 1]$, e.g. by the min-max normalization.

Due to the data availability and the data quality, only three characteristics are considered in above definition. More information, like the length of the time that each customer spend on one item and the customer performances (e.g. regret or not) after the buying decision, should be helpful for better measuring indecisiveness. Meanwhile, it is necessary to make a distinction between the customer behaviors of surfing and shopping. However, we should note that the following processes of indecisiveness mining are general enough to deal with any specific measure of observed indecisiveness.

C. Indecisive Index Learning

We learn latent indecisive indexes for both customers and items by a probabilistic factor-based model.

It is straightforward and quite necessary for us to figure out the reasons behind the observed indecisiveness in each session. We think there are generally two types of unobserved reasons: the customer-specific factor and the items-specific factor, and we use "Indecisive Index" to represent these factors. That is, the customer with high indecisive index (i.e. she is a kind of person that usually has difficulties in decision-making) will lead to a high level of observed indecisiveness, and the bundle of items with a high indecisive index (i.e. it is hard to make choice from these items, e.g. buy an Iphone or a Sumsung Note?) will also lead to high observed indecisiveness, and vice versa. Here, we use Y_i and Z^S_j to represent the indecisive index for customer U_i and session S_j, respectively. Indeed, the indecisive index of one session is reflected by the set of the item bundles in this session (Definition 1). These item bundles for S_j is denoted by $S_B(j) = \{B_{j(p_1)}, B_{j(p_2)}, ..., B_{j(p_T)}\}$, e.g. for S_4 in Table II, $S_B(4)$ is $\{\{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$[5]. In other words, the difficulty of making choice from item a, b, c (or the comparisons among these items) leads to the generation of current session. Thus, the latent indecisive index of session S_j is dependent on the latent indecisive index of all the item bundles in $S_B(j)$. We formulate this influence as follows:

$$Z^S_j = \sum_{t=1}^{T} w_{j(p_t)} Z^B_{j(p_t)}, \tag{2}$$

where $Z^B_{j(p_t)}$ is the latent indecisive index of the t-th item bundle, i.e. $B_{j(p_t)}$, in $S_B(j)$. $w_{j(p_t)}$ is the weight of this item bundle, and its value could be determined based on the number of appearance of the items in this bundle. For instance, the items in $\{b, c\}$ and $\{a, b, c\}$ appear 2 times and 4 times in session S_4 in Table II. We normalize all the weights so that $\sum_{t=1}^{T} w_{j(p_t)}=1$. As said in Section III-A, we have removed the cold-start customers and item bundles in preprocessing.

The above equation indicates that the estimation of the latent indecisive index of a session Z^S_j is the weighted average of the latent indecisive indexes of its frequent item bundles $Z^B_{j(p_t)}$. As we try to learn the latent Y_i and $Z^B_{j(p_t)}$ so as to precisely estimate D^S_{ij}, we could use a probabilistic linear model with Gaussian observation noise. The graphical model is shown in Figure 2. Thus, the conditional probability of the observed indecisiveness D^S of all the sessions is

$$p(D^S|Y, Z^B, W_B, \sigma^2) = \prod_{i=1}^{M}\prod_{j=1}^{N}[N(D^S_{ij}|g(Y_i Z^S_j), \sigma^2)]^{H_{ij}}, \tag{3}$$

where $N(D^S_{ij}|g(Y_i Z^S_j), \sigma^2)$ is the probability density function of the Gaussian distribution with mean $g(Y_i Z^S_j)$ and variance σ^2, and H_{ij} is the indicator function that is equal to 1 if customer U_i has session S_j and equal to 0 otherwise. We should note that, we represent D^S and H in the format of a 2-dimensional matrix (i.e. D^S_{ij} and H_{ij}) for better illustration. Indeed, different from traditional graphical models, there is usually only one none-zero entry in each column (stands for one session) of the matrix, since the customers are unique for each session. However, this way of representation will have no impact on our implementation, as we actually just consider the frequent item bundles instead of the sessions.

We set the function $g(x)$ as $g(x) = (1 - e^{-x})/(1 + e^{-x})$ and $x > 0$, which bounds $Y_i Z^S_j = Y_i(\sum_{t=1}^{T} w_{j(p_t)} Z^B_{j(p_t)})$ within the range of $[0, 1)$. We also place zero-mean Gaussian priors on the indecisive indexes of customers and item bundles. Hence, by a Bayesian inference we could learn the posterior probability of the latent indecisive indexes Y and Z^B as follows:

$$\begin{aligned} &p(Y, Z^B|D^S, W_B, \sigma_U^2, \sigma_B^2, \sigma^2) \\ &= \textstyle\prod_{i=1}^{M}\prod_{j=1}^{N}[N(D^S_{ij}|g(Y_i(\sum_{t=1}^{T} w_{j(p_t)} Z^B_{j(p_t)})), \sigma^2)]^{H_{ij}} \\ &\quad \textstyle\prod_{i=1}^{M}[N(Y_i|0, \sigma_U^2)]\prod_{j=1}^{P}[N(Z^B_j|0, \sigma_B^2)]. \end{aligned} \tag{4}$$

Y and Z^B can be learned by maximizing this posterior or log-posterior, and this is equivalent to minimizing the sum-of-squared-errors objective function with respect to Y and Z^B:

$$E = \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{N} H_{ij}[D^S_{ij} - g(Y_i(\sum_{B_k \in S_B(j)} w_{jk} Z^B_k))]^2 + \frac{\lambda_U}{2}\sum_{i=1}^{M} Y_i^2 + \frac{\lambda_B}{2}\sum_{k=1}^{P} Z^{B^2}_k, \tag{5}$$

where $\lambda_U = \sigma^2/\sigma_U^2$ and $\lambda_B = \sigma^2/\sigma_B^2$. Note that, B_k is used in Equation (5) to represent one item bundle rather than $B_{j(p_t)}$ in Equations (2), (3), and (4). Actually, these two types of notations have the same meanings, where B_k is one specific item bundle in terms of the entire set of item bundles and $B_{j(p_t)}$ is an item bundle in session S_j, i.e. for each $B_{j(p_t)}$ there is one and only one B_k, which makes $B_{j(p_t)}=B_k$ (similarly, $w_{j(p_t)} = w_{jk}$). As they could represent the item bundles for different scenarios, in this paper, we use these two notations simultaneously. A local minimum of the objective function given by Equation (5) can be obtained by performing gradient descent in Y_i and Z^B_k. For simplicity, we set $\lambda_U = \lambda_B$ for the experiments, and the initial values of Y_i and Z^B_k are randomly

[5]To make the distinction between items and item bundles more significantly, in current work, we do not consider the 1-item bundle, e.g., $\{a\}$, $\{b\}$, $\{c\}$.

sampled according to Gaussian distribution. However, during training, we forcibly set Y_i and Z_k^B equal to 0 when they are smaller than 0, making the values stay in the range of $[0, +\infty)$ so as to be comparable to each other. Eventually, Y_i and Z_k^B can be learned to express the indecisive index of customer U_i and item bundle B_k, respectively.

We call our probabilistic matrix factorization model for indecisiveness learning as IMF. Actually, as a type of probabilistic model [31], we found the graphical representation of IMF looks similar to several of the existing models, i.e. RSTE [32] and SocialMF [33]. However, IMF is still quite different from these traditional ones. For instance, in IMF there are three types of latent factors and one of them, i.e. the latent factor for each session (Z_j^S), is determined by its internal subsections (sub-factors, i.e. $B_{j(p_t)}$); while for both RSTE and SocialMF (which are used for rating prediction in recommender systems), there are two types of latent factors and one of them, i.e. the latent factor for users, are influenced by the same type of latent factors from external friends. Thus, the meaning underlying the inference processes is significantly different and these models are suitable for different scenarios.

At last, we should note that, for better explanation, we have deliberately simplified our model. For instance, we only treat the indecisive index as a scalar. Actually, it can be also viewed as a vector, and each entry of which may store the scale of indecisiveness from the customer to one aspect of the item bundle (e.g. price). Meanwhile, we currently focus on the item bundles with no less than 2 items in each bundle, but our model also works for the one-item bundles.

D. Applications

Many applications can be derived, for instance, the analysis of indecisiveness will be helpful for the better understanding of customers [17], [22]. Since some of the customer profiles (e.g. age and gender) are not available, we mainly focus on two of the examples including the competitive item detection, and more importantly, the item bundle recommendation.

Competitive Items Detection. In some of the existing marketing solutions, the items that are frequently bought together by the customers will be combined together for better selling [34]. However, in the customers' behavior (e.g. click) sessions, it is more often that the items appearing together may compete against each other. In other words, the indecisiveness analysis proposed in this paper provides a promising direction for the online retailers to figure out the competitive items. Along this line, we could first run IMF model to get the frequent item bundles and their indecisive indexes. Then, suppose item a is a product of the given retailer, the items that are in the same bundles (especially the bundles having very large indecisive indexes) with a could be identified as the competitive items of item a. In this way, the competitive items can be figured out from a data-driven perspective, and the results help retailers conduct better item assortment and marketing strategy.

Item Bundle Recommendation. Recommender systems target on recommending the right items to the right customers at anytime and anywhere [4]. However, traditional

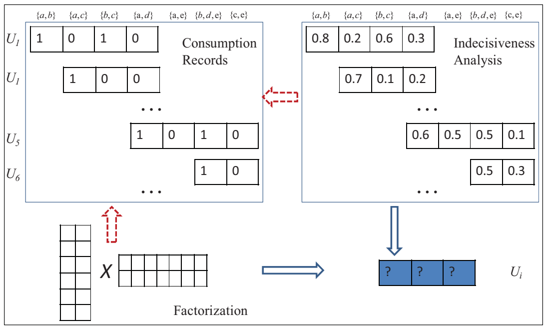


Fig. 3. The flowchart of Item Bundle Recommendation.

recommendation algorithms seem to be helpless when the customers already suffer from the indecisiveness. First, if the recommended items are similar to the ones that this customer is currently viewing, she will be more frustrated since this will make the decision even harder. Second, if the recommended items has nothing to do with the ones that this customer is currently viewing, she will be of no interest since the recommendations are beyond her current concerns.

Nevertheless, the recommendation service is still needed when the customers are suffering from the indecisiveness. Suppose the given customer is currently viewing items $\{a,b,c,d,e\}$, and if we could predict that the final consumption (buying action) will possibly happen in bundle $\{a, b\}$ (i.e. either a or b), then we can only recommend this item bundle during the customer's indecisiveness, and thus, save her from the useless actions. From the customers' perspective, they get more powerful extra signs to help them make a final buying decision. From the service providers' perspective, they could lead the customers to the "right" choice, i.e. by providing choice filtering, the indecisiveness of the customers may be alleviated and the customers will perform a buying behavior with a higher probability (We will show this experimentally in Section IV-A). Thus, the service providers could get more profit. Unfortunately, based on the above analysis, the traditional recommendation algorithms are not suitable for this task. To this end, in this paper, we propose an idea of item bundle recommendation for consumption inducement, i.e. when the customers are suffering from indecisiveness, we could recommend an item bundle from which she has the highest probability to choose (consume). For each indecisive session S_j of customer U_i, the candidate item bundles are chosen from $S_B(j)$, and thus these item bundles are usually with smaller sizes. Besides, this item bundle recommendation could provide extra signs to help support the customers' preference, and thus, making the consumption much easier.

In the following, we propose a specific recommendation algorithm to show the way of incorporating customer indecisiveness to conduct item bundle recommendation. Similarly, indecisiveness could also be exploited to help enhance the performance of other algorithms (e.g. the Neighbor-based collaborative filtering algorithms [4] as shown experimentally) based on the following assumption.

Assumption *Given a customer and one of her behavior session consisting of several item bundles, the higher inde-*

cisiveness of one item bundle, the higher probability that the finally consumed item locates in this bundle.

In other words, there is a positive correlation between the consumption probability and the indecisiveness value. Since the customer feedback is implicit (either 1 or 0, i.e. buy the item(s) in this bundle or not), we borrow the idea from Bayeisan Personalized Ranking (BPR) [35] and propose a novel algorithm to incorporate indecisiveness for Item Bundle Recommendation (IBR). The flowchart is shown in Figure 3. Specifically, given the consumption records and the indecisiveness analysis results of each session (Section III-C), we hope to get two low-rank matrices (in terms of customers and item bundles, respectively) using matrix factorization. Here, the optimization objective is that combining the factorized low-rank matrices and the given indecisiveness analysis results could well fit the partial ranking orders (based on consumption or not, and we will illustrate the details later) existing in the consumption records (shown by the dashed arrows in red in Figure 3). Then, using these low-rank matrices and the indecisiveness analysis results we can predict the final consumption behavior of one testing session, e.g. the ranking orders (the consumption behaviors) from U_i to bundles $\{b, c\}, \{a, d\}$ and $\{a, e\}$ (shown by the arrows in blue in Figure 3). In the consumption records, one row of records stand for the finally consumed item in a session: if the consumed item locates in this bundle, the corresponding value of this column is 1, otherwise 0. For instance, in the first session of U_1, the finally consumed item locates in bundle $\{a, b\}$ and $\{b, c\}$, i.e. item b is actually the choice of the customer. In the indecisiveness analysis part, one row of records represent the computed indecisiveness from each customer to the item bundles in this session (based on IMF model, and the exact way of computation will be shown lately). For instance, 0.8 in the first entry means U_1 shows 0.8 indecisiveness to item bundle $\{a, b\}$ in this session. Thus, according to Assumption 1, the final consumption of U_1 may (does) locate in $\{a, b\}$.

We formulate the prediction of the buying probability from customer u to the i-th item bundle in a session s as:

$$\hat{r}^s_{ui} = (Q'_u \cdot V_i) + \alpha_u (Y'_u \cdot w_{si} \cdot Z^B_{si}), \tag{6}$$

where Q_u and V_i are the factorized low-rank feature vectors for customer and the item bundle respectively, and Y_u, w_{si} and Z^B_{si} are the same to those in IMF model. Personalized parameter α_u balances the contributions of the indecisiveness. Suppose the right part is given (i.e. $(Y'_u \cdot w_{si} \cdot Z^B_{si})$), we would like to get Q_u, V_i and α_u for estimating the observed consumption records. More specifically, as the consumption records are implicit feedbacks, we resort to the learning to rank strategy [35] and estimate the personalized (sessonalized) ranking order of two item bundles. This estimator $\hat{r}^s_{uij}$ could be defined as $\hat{r}^s_{uij} = \hat{r}^s_{ui} - \hat{r}^s_{uj}$. If using $i >_s j$ to denote the ranking order of two item bundles, we could define the individual probability that in session s the customer prefers the i-th item bundle to the j-th item bundle as $p(i >_s j | Q, V, \alpha_u) = \sigma(\hat{r}^s_{uij})$, where $\sigma(x)$ is the logistic sigmoid $1/(1+e^{-x})$. Further, if using $>_s$ to denote all the ranking orders, the posterior probability should be maximized is

$$p(Q, V, \alpha | >_s) \propto p(>_s | Q, V, \alpha) p(Q, V, \alpha). \tag{7}$$

TABLE IV
THE STATISTICS OF THE DATASET BEFORE AND AFTER PRUNING.

	Original Data	Pruned Data
#Customers	9,774,184	1,998,112
#Items	8,133,507	50,700
#Sessions	234,496,841	4,182,243
#Actions	1,333,729,303	44,000,123
#Item Bundles	\	141,951

Here, we could also introduce Gaussian priors on factorized feature vectors Q, V and α, respectively. These values could be learned by minimizing the following objective function:

$$L = \sum_{s,u,i,j} (\lambda_\theta ||Q_u||^2 + \lambda_\theta ||V_i||^2 + \lambda_\theta ||V_j||^2 + \lambda_\theta ||\alpha_u||^2 - ln\sigma(\hat{r}^s_{ui} - \hat{r}^s_{uj})), \tag{8}$$

where λ_θ stands for the model-specific regularization parameters. Then we perform gradient descent on Q, V and α to get the local minimum of the above objective function. Due to the space limitation, we only show the gradient in α:

$$\frac{\partial L}{\partial \alpha_u} = \frac{-[(Y'_u \cdot w_{si} \cdot Z^B_{si}) - (Y'_u \cdot w_{sj} \cdot Z^B_{sj})]}{1 + e^{Q_u \cdot (V_i - V_j) + \alpha_u [(Y'_u \cdot w_{si} \cdot Z^B_{si}) - (Y'_u \cdot w_{sj} \cdot Z^B_{sj})]}} + \lambda_\theta \alpha_u. \tag{9}$$

After getting Q, V and α from this model, we could apply them to estimate the preference of the customers in the testing set, and thus, the item bundles can be recommended based on the estimated preference.

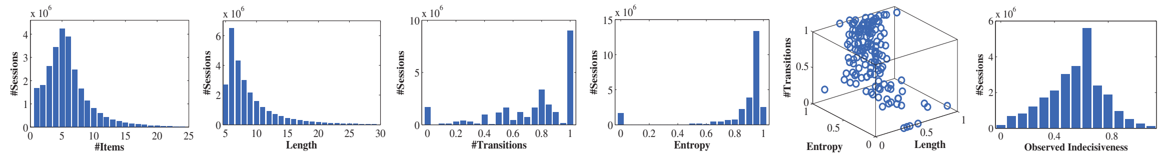
IV. EXPERIMENTS

The dataset *Rec-Tmall*[6], that contains a large-scale behavioral logs of online customers, is stored in a distributed platform, call *TianChi*, for open research. Both Rec-Tmall and TianChi are provided by Tmall (a business unit of Alibaba), one of the largest B2C online retail platform in China.

Dataset. To make sure the reliability of the experimental results, we conduct some preprocessing on the dataset. First, we split the records into sessions based on a simple but widely used rule [29]: two actions are split into different sessions if the time interval between them exceeds a threshold, e.g. 1 hour. Then, the cold-start customers (only show up once) and the spider candidates are removed. Here, we treat the customers with more than 1000 actions in one session as the spiders and remove them. We get the frequent item bundles, that appear in more than 100 sessions, by the Apriori-like algorithm [30]. The statistics of the dataset before and after preprocessing could be found in Table IV, where we can see that a large number of the items are removed due to the long tail effect. The dataset was then divided into a training set and a test set, where 80% records of each customer were randomly chosen to be part of the training set and the remaining records were used for testing.

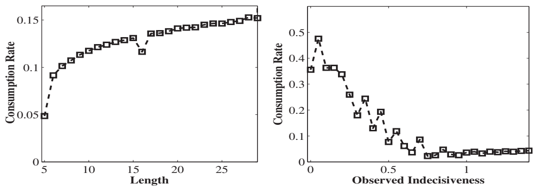
Experimental Platform. The experimental platform, TianChi, is running on an Open Data Processing Service (ODPS), which is developed by Alibaba for dealing with big data, e.g. in the scale of TB/PB. TianChi supplies some interfaces and operations for the MapReduce programming model. Due to the privacy issue and space limitation, details of the configurations of each worker machine and the distributed implementation of each model will be omitted.

[6] http://tianchi.aliyun.com/datalab/dataSet.htm?spm=5176.100073.888.5.HraQ79&id=2



(a) #Items Distr. (b) #Actions Distr. (c) #Trans. Distr. (d) Entropy Distr. (e) 3D Scatter (f) Inde. Distr.

Fig. 4. Some statistics from the observed records.



(a) Length *vs* Rate (b) Indecisiveness *vs* Rate

Fig. 5. The statistics on the sessions with consumption.

A. Existence of Indecisiveness

We illustrate some statistics from the records, and they provide the evidence of the existence of indecisiveness.

First, sub-figures in Figure 4 show (a) the distribution of the number of items in each session, (b) the distribution of the number of actions (i.e. session length), (c) the distribution of the number of transitions, (d) the distribution of the entropy of items, (e) the 3D scatter of hundreds of randomly selected sessions, and (f) the distribution of the observed indecisiveness values, respectively. From Figure 4(a-d) we could observe that the customer in each session usually has actions (the average length is 11) on multiple items (an average of 8 items), and these actions often transfer among the items with a high entropy. Combining the results in Figure 4(e), $Length(\cdot)$, $Trans(\cdot)$ or $Entropy(\cdot)$ captures one aspect of the session and their values are usually different from each other, i.e., it is necessary to consider all of them for measuring the indecisiveness of one session (We will also support this claim in the later experiments). Based on these observations and without loss of generality, in both Figure 4(f) and the following experiments, we let $F(Length(S_j), Entropy(S_j), Trans(S_j))$ in Definition 1 equal to $Length(S_j) * Entropy(S_j) * Trans(S_j)$ for measuring the observed indecisiveness. The reasonability of this measurement will be evaluated in the following subsections. Here, we only show the statistical results in Figure 4(f) (a higher value means a higher level of indecisiveness). In summary, Figure 4 shows that the customer has to make so many comparisons in each session.

Then, We only focus on the sessions with consumption behaviors (buying actions) and some results on these sessions are shown in Figure 5. Here, we try to figure out the relationship between the length of the sessions ($Length(\cdot)$) and the final consumption rate (i.e. the number of buying sessions divided by the total sessions at that length). Actually, we find that the longer the session the higher probability that there will be a buying behavior, which implies that the customers are usually very careful to make a consumption. In Figure 5(b), we illustrate the correlation between the observed indecisiveness and the consumption rate, where we can find that the higher indecisiveness the lower rate of consumption (Here, we also suppose the measurement of our observed indecisiveness is reasonable). In other words, when the customer feels very

TABLE V

USER STUDY RESULTS.

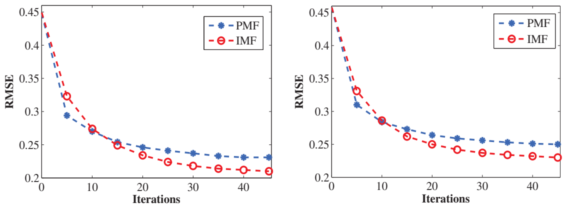
Metrics		$Length(\cdot)$	$Trans(\cdot)$	$Entropy(\cdot)$	Indecisiveness
Jaccard Similarity	Mean	0.736	0.372	0.063	**0.849**
	$\|z\|$	4.091	20.791	33.095	\
Euclidean Distance	Mean	1.100	2.033	2.829	**0.596**
	$\|z\|$	5.603	19.035	28.739	\

indecisive she will not make a buying choice. Indeed, $Length(\cdot)$ measures whether the customer is serious about this consumption while indecisiveness measures the hardness of the choice, and the difference between these two measurements could be clearly concluded from the different correlation results in Figure 5(a) and Figure 5(b). Thus, Figure 5 further implies that the service providers should induce the customer to have longer considerations, and meanwhile, it is necessary to make item bundle recommendations to alleviate the indecisiveness.

B. Indecisiveness Analysis

We experimentally illustrate the reasonability of our measurement on observed indecisiveness. Then, we evaluate the predictive power of IMF for latent indecisive indexes learning.

We still let $F(Length(S_j), Entropy(S_j), Trans(S_j))$ in Definition 1 equal to $Length(S_j) * Entropy(S_j) * Trans(S_j)$. Indeed, it is not easy to directly evaluate the reasonability of the measurement, since even a clear definition of indecisiveness is lacking [10]. As an alternative, the performance of the measurement could be further proved by the applications, e.g. the item bundles recommendation, which will be shown lately. Nevertheless, we also refer to the human judgement as a metric. Since it is impossible for us to contact the customers in the dataset directly, we conduct another type of user study. Specifically, we first choose a number of behavior session pairs from the dataset, and then publish an online questionnaire in diaochapai.com, where each pair of behavior sessions are shown to the volunteers. Then, we ask the volunteers to choose the session that seems to be more indecisive of the customer (the binary value 1 represents the first session and 0 represents the second session of this session pair, respectively). In total, we collected 954 qualified choices from 106 volunteers to 9 pairs of behavior sessions (most of these volunteers have a Master or Ph.D degree in business, management or computer science, and the careless answers that are given quickly, e.g. in less than 5 seconds, are removed). Next, we compare the similarity/distance between the choice vector (the value of each entry, i.e. 1 or 0, records the choice of volunteer to one session pair) of the volunteers and the results computed based on $Length(\cdot)$, $Entropy(\cdot)$, $Trans(\cdot)$, and our indecisiveness measurement (Indecisiveness for short). Finally, the average value (Mean) under Jaccard similarity and Euclidean distance are shown in Table V, where we also provide the z-test results ($|z|$) between Indecisiveness and others. From this table we can observe that the choice of our Indecisiveness is the most similar to the choice of the volunteers, and the differences between the similarities/distances obtained by Indecisiveness



(a) Training Set (b) Test Set

Fig. 6. The results of indecisive index learning.

and the other measurements are statistically significant with $|z| \geq 1.96$ and thus $p \leq 0.05$. Indeed, we have also evaluated this user study under other metrics, e.g., Cosine similarity, Accuracy and Hamming distance, and the similar observations are omitted due to the space limitation.

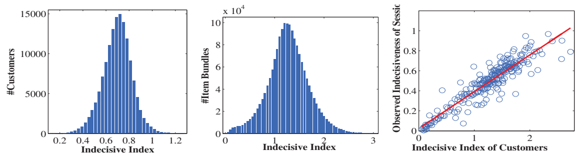
Given the observed indecisiveness, we then evaluate the predictive power of IMF model. Since IMF is also a type of probabilistic model, we compare it with the state-of-the-art Probabilistic Matrix Factorization (PMF) model [31]. In PMF, the observed indecisiveness D^S_{ij} is estimated for multiple times, and each time it is directly computed by Y_i and $Z^B_{j(p_t)}$ rather than Y_i and Z^S_j in IMF. Thus, the major difference between IMF and PMF is considering the item bundles in one session as a whole or not. We measure the effectiveness of these models by the metric of Root Mean Square Error (RMSE). We should note that we also tried some other prediction methods, such as Singular Value Decomposition (SVD). As our data is extremely sparse (with none-zero rate 0.007%), none of these methods could get a comparable performance. For comparison, we set the parameters equally, e.g. setting 0.01 for all the λs. The final performances of them on the training set and test set are shown in Figure 6, where we can see that IMF is much more effective than PMF on both two sets (with nearly 10% of the prediction improvement after convergence). We also record the running time of these two methods, and we observe that the average time per iteration for both of them is about 270 seconds, when we use 20 worker machines in the cluster.

Next, we illustrate the distribution of the customer indecisive indexes, the item bundle indecisive indexes and the correlations between the latent indecisive indexes (we take the customer index as an example) and the observed indecisiveness of sessions in Figure 7. From Figure 7(a) and Figure 7(b) we can see that the learnt indecisive indexes also follow the Gaussian distribution, and from Figure 7(c) we find that the latent indecisive index of customers are positively correlated with the observed indecisiveness of the behavior sessions. These results also imply the effectiveness of the IMF model.

C. Item Bundle Recommendation

We present the performance comparison between IBR and the benchmarks on the sessions with consumptions (buying actions). The results also provide convincing conclusions on the reasonability of our measurement on observed indecisiveness.

First, we prove Assumption 1, i.e. indecisiveness helps item bundle recommendation. To this end, in Figure 8 we illustrate the distribution of the location of the consumed items in each buying session. Here, the horizontal axis are the rank locations (in percentage, i.e. the smaller location the higher indecisiveness $(Y'_u \cdot w_{si} \cdot Z^B_{si})$) of the consumed items and the



(a) Customers (b) Item Bundles (c) Correlations

Fig. 7. Statistical results of indecisive indexes.

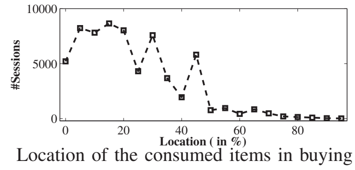


Fig. 8. Location of the consumed items in buying sessions.

vertical axis store the number of sessions. Generally, the smaller location the higher consumption probability, i.e. there is a positive correlation between the consumption probability and the indecisiveness. Thus, Assumption 1 holds and we should exploit indecisiveness for better item bundle recommendation.

Second, we evaluate the performance of our item bundle recommendation algorithm IBR. Here, we choose six methods ,i.e., REC, FRE, IND, BPR, KNN and IKNN as baselines. Among them, REC and FRE are the straightforward solutions, i.e. recommending the item bundles based on the location of their last appearances (REC, "recency") or the total number of appearances (FRE, "frequency") in the session. Further, IND is the method that directly adopts the learned indecisiveness values for ranking each item bundle (i.e. $(Y'_u \cdot w_{si} \cdot Z^B_{si})$), while BPR is the bayeisan personalized ranking method without considering the impact of indecisiveness [35] (i.e. $\hat{r}^s_{ui} = (Q'_u \cdot V_i)$). Thus, they could be viewed as the preliminary versions of our IBR. Since the customer feedback is implicit rather than the explicit ratings, we choose another ranking-oriented and neighbor-based collaborative filtering method [4], KNN, where we first find the K-nearest sessions by Jaccard similarity (based on the item bundles) for each session, and then the most frequently bought item bundles in these sessions are recommended to the target customer in current session. Finally, IKNN could be viewed as the weighted version of KNN, where the indecisiveness of each item bundle is treated as weight when finding the nearest sessions. For the purpose of evaluation, we adopted AUC (Area Under the ROC Curve) and Top-1 as metrics, where AUC measures the percentage of the item bundle pairs ranked in the correct order and Top-1 indicates the precision of the recommended top one item bundles [4]. We record the best performance of each algorithm by tuning their parameters and also give their running time.

The final results are shown in Table VI. In terms of effectiveness, IBR, which combines indecisiveness and factorization to collaboratively analyze a group of customers, outperforms all the baselines for both AUC and Top-1, and this also shows the rationality of our measurement on indecisiveness. By incorporating the impact of indecisiveness, IKNN also performs much better than the preliminary version, KNN. Since all the item bundles in one session meet the customer interests, KNN can not make a good distinction in such a situation. Meanwhile, both IND and FRE can be categorized into the frequency based methods, and they perform similarly.

TABLE VI
THE RESULTS OF ITEM BUNDLE RECOMMENDATION.

Metric \ Method	REC	FRE	KNN	IKNN	IND	BPR	IBR
AUC	0.757	0.802	0.521	0.615	0.818	0.826	**0.893**
Top-1	0.762	0.869	0.288	0.406	0.895	0.647	**0.924**
RunTime(Sec.)	**100**	102	130	135	102	110	113

Actually, we observed BPR could perform as well as IBR in the training set, however, it becomes much worse in the test set, i.e. suffering from overfitting. In terms of efficiency, as it is hard to capture the exact running time of each algorithm in a platform with workloads changing from time to time, we only record the average time. Indeed, there is no significant difference, and REC is slightly better than others.

Third, we illustrate the distribution of α_u, which balances the impact of the indecisiveness on each customer (Figure 9). We find it could also be fitted by a Gaussian distribution. Meanwhile, most of the α_us are far from 0, i.e. the indecisiveness contributes to the final item bundle recommendation.

D. Case Study

We conduct a case study to explore another application of our indecisiveness mining, i.e. competitive items detection, and we show both global competitions and personal competitions.

In the global competition detection, we first rank the item bundles based on their indecisive indexes and then we manually select three typical bundles from different categories, i.e., Furnish, Jewelry and Mobile Phone. As shown in Figure 10, the items in these three bundles (2-item bundles) look similar and also have some other properties in same. Let's take the two mobile phones (Figures 10(e) and (f)) as an example. These two phones are HuaWei C8813[7] and CoolPad 8720[8] from different retailers, with the similar prices (less than 700 RMB), Operating System (Android), hardware configurations, and etc. Thus, they attract the customers from the same segment and should be a strong competitor to each other.

For the personalized competition detection, we first select two items from the categories of Skirt and Couch as shown in Figure 11(a) and Figure 11(d). Then, for each given item, we find the item bundles containing this item and also having high indecisive indexes. Thus, other items in these item bundles are the personalized competitors of this given one. The URLs of these items are omitted due to the limited space. Here, we show two of these competitors for each given item. We can see that the competitors look similar to the given one and are hard for customers to make the consumption decision. In the real-world applications, the indecisive indexes could be explicitly given to the service providers for measuring this competition.

Finally, we should note that these competitive items may be also mined by simpler solutions, e.g. based on the frequent item bundles where exactly (with high probability) one of the items were bought. However, there are several limitations of this solution: It could not include the impact from the customers (e.g. different indecisive indexes of customers), and it cannot quantify the degree of the competition very well.

[7]http://detail.tmall.com/item.htm?spm=a220o.1000855.1000983.1.faab2D&id=38866802884&standard=1

[8]http://detail.tmall.com/item.htm?spm=a220o.1000855.1000983.1.M1KEOq&id=18101417851&standard=1

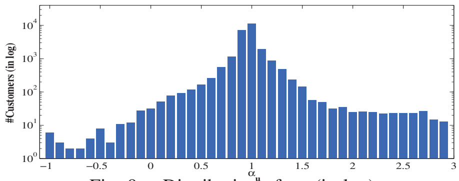


Fig. 9. Distribution of α_u (in log).

V. DISCUSSION

From the experimental results, we can see that our solutions work well for automatically mining indecisiveness from massive customer behaviors. Specifically, the measurement on observed indecisiveness agrees with the common sense assessment, and IMF model could capture the latent reasons for indecisiveness. Furthermore, by considering indecisiveness, our IBR algorithm could more accurately locate the item bundles from which the customer will finally consume.

As a general and flexible framework, some components of our indecisiveness mining can be extended for specific applications: First, in this paper, we only use limited information to measure the observed indecisiveness from customer behaviors, and we believe much more characteristics (e.g. the length of the time that each customer spend on one item) could be considered. Meanwhile, more reasonable ways of combining these information for indecisiveness measurement should also be designed, e.g. considering other functions than $log(\cdot)$ for $Length(S_j)$ or incorporating the metrics from other research domains such as psychology; Second, for better explanation, we treat indecisive index as a simple scalar, and it could also be a vector of latent factors. At the same time, we only focus on the item bundles with no less than 2 items in each bundle, and it is of great interest to find out the indecisive indexes of each single item; Third, this data-driven way of mining indecisiveness could be helpful to many more applications. Indeed, given that the topic of mining indecisiveness from massive customer behaviors has largely been neglected, there are too many research directions remain to be explored.

VI. CONCLUSION

In this paper, we provided a focused study on the problem of automatically mining indecisiveness from massive customer behaviors. Along this line, we first gave a general definition to measure the scale of the observed indecisiveness in each behavior session without manual interventions. Then, by fitting the observed indecisiveness, we proposed a probabilistic factor-based model which could learn the latent indecisive indexes for both customers and item bundles. Next, considering the impact of indecisiveness, we designed a recommendation algorithm to better locate the item bundles that one customer may choose(consume) from. Finally, we conducted extensive experiments on large-scale behavioral logs of online customers in a distributed environment using MapReduce. The experimental results clearly demonstrated the effectiveness of our proposed models. Actually, given any specific measure of observed indecisiveness, the proposed framework is general enough to address the problem of indecisiveness mining. Thus, we hope this study could lead to more future work.

(a) Children bed-1 (b) Children bed-2 (c) Bracelet-1 (d) Bracelet-2 (e) Mobile phone-1 (f) Mobile phone-2

Fig. 10. Three pairs of typical items that are competitive to each other.

(a) Given Skirt (b) Competitor-1 (c) Competitor-2 (d) Given Couch (e) Competitor-1 (f) Competitor-2

Fig. 11. The competitors of one skirt and one couch, respectively.

Acknowledgements. The authors thank Alibaba for providing the dataset and the experimental environment, and thank Dr. John Krumm for allowing us using his social information, and thank Guowei Ma and Le Wu for the valuable discussions and suggestions. This research was partially supported by grants from the National Science Foundation for Distinguished Young Scholars of China (Grant No. 61325010), the Natural Science Foundation of China (Grant No. 61403358), the Anhui Provincial Natural Science Foundation (Grant No. 1408085QF110) and the CCF-Tencent Open Research Fund. Qi Liu gratefully acknowledges the support of the Youth Innovation Promotion Association, CAS.

References

[1] B. Fu, J. Lin, L. Li, C. Faloutsos, J. Hong, and N. Sadeh, "Why people hate your app: Making sense of user feedback in a mobile app store," in *ACM SIGKDD*. ACM, 2013, pp. 1276–1284.

[2] R. Li, S. Wang, H. Deng, R. Wang, and K. C.-C. Chang, "Towards social user profiling: unified and discriminative influence model for inferring home locations," in *ACM SIGKDD*. ACM, 2012, pp. 1023–1031.

[3] A. Chernev, "Product assortment and consumer choice: An interdisciplinary review," *Found Trends Mark*, vol. 6, pp. 1–61, 2011.

[4] P. B. Kantor, L. Rokach, F. Ricci, and B. Shapira, *Recommender systems handbook*. Springer, 2011.

[5] A. Anderson, D. Huttenlocher, J. Kleinberg, and J. Leskovec, "Steering user behavior with badges," in *WWW*, 2013, pp. 95–106.

[6] A. Bechara and A. R. Damasio, "The somatic marker hypothesis: A neural theory of economic decision," *Games and economic behavior*, vol. 52, no. 2, pp. 336–372, 2005.

[7] E. Polman, "Effects of self–other decision making on regulatory focus and choice overload." *Journal of personality and social psychology*, vol. 102, no. 5, p. 980, 2012.

[8] E. Rassin and P. Muris, "Indecisiveness and the interpretation of ambiguous situations," *Personality and Individual Differences*, vol. 39, no. 7, pp. 1285–1291, 2005.

[9] E. Rassin, P. Muris, I. Franken, M. Smit, and M. Wong, "Measuring general indecisiveness," *Journal of Psychopathology and Behavioral Assessment*, vol. 29, no. 1, pp. 60–67, 2007.

[10] E. Rassin, "A psychological theory of indecisiveness," *Netherlands journal of psychology*, vol. 63, no. 1, pp. 1–11, 2007.

[11] R. O. Frost and D. L. Shows, "The nature and measurement of compulsive indecisiveness," *Behaviour Research and Therapy*, vol. 31, no. 7, pp. 683–IN2, 1993.

[12] V. Germeijs and P. De Boeck, "A measurement scale for indecisiveness and its relationship to career indecision and other types of indecision." *European Journal of Psychological Assessment*, vol. 18(2),113, 2002.

[13] B. Scheibehenne, R. Greifeneder, and P. M. Todd, "Can there ever be too many options? a meta-analytic review of choice overload," *Journal of Consumer Research*, vol. 37, no. 3, pp. 409–425, 2010.

[14] E. K. Cho, U. Khan, and R. Dhar, "Comparing apples-to-apples or apples-to-oranges: Choice difficulty and home buying," *Keller Center Research Report*.

[15] A. Sela, J. Berger, and W. Liu, "Variety, vice, and virtue: How assortment size influences option choice," *Journal of Consumer Research*, vol. 35, no. 6, pp. 941–951, 2009.

[16] A. Chernev, U. Böckenholt, and J. Goodman, "Choice overload: A conceptual review and meta-analysis," *Journal of Consumer Psychology*, 2014.

[17] M. E. Brooks, "Management indecision," *Management Decision*, vol. 49, no. 5, pp. 683–693, 2011.

[18] M. C. Willemsen, B. P. Knijnenburg, and et al., "Using latent features diversification to reduce choice difficulty in recommendation lists," *ACM RecSys*, vol. 11, pp. 14–20.

[19] J. Goodman, S. Broniarczyk, J. Griffin, and L. McAlister, "Help or hinder? when recommendation signage expands consideration sets and heightens decision difficulty," *Journal of Consumer Psychology*, vol. 23, no. 2, pp. 165–174, 2013.

[20] S. Botti and A. L. McGill, "The locus of choice: Personal causality and satisfaction with hedonic and utilitarian decisions," *Journal of Consumer Research*, vol. 37, no. 6, pp. 1065–1078, 2011.

[21] S. M. Broniarczyk and J. Griffin, "Decision difficulty in the age of consumer empowerment," *Journal of Consumer Psychology*, 2014.

[22] E. Rassin and P. Muris, "To be or not to be indecisive: Gender differences, correlations with obsessive–compulsive complaints, and behavioural manifestation," *Personality and Individual Differences*, vol. 38, no. 5, pp. 1175–1181, 2005.

[23] R. Dhar and M. Gorlin, "A dual-system framework to understand preference construction processes in choice," *Journal of Consumer Psychology*, vol. 23, no. 4, pp. 528–542, 2013.

[24] F. Zhang, N. J. Yuan, D. Lian, and X. Xie, "Mining novelty-seeking trait across heterogeneous domains," in *WWW*, 2014, pp. 373–384.

[25] Q. Liu, E. Chen, H. Xiong, C. H. Ding, and J. Chen, "Enhancing collaborative filtering by user interest expansion via personalized ranking," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 42, no. 1, pp. 218–233, 2012.

[26] Y. C. Zhang, D. Ó. Séaghdha, and et al., "Auralist: introducing serendipity into music recommendation," in *ACM WSDM*, 2012, pp. 13–22.

[27] H. Zhu, H. Xiong, Y. Ge, and E. Chen, "Mobile app recommendations with security and privacy awareness," in *ACM SIGKDD*. ACM, 2014, pp. 951–960.

[28] P. Yin, P. Luo, W.-C. Lee, and M. Wang, "Silence is also evidence: interpreting dwell time for recommendation from psychological perspective," in *ACM SIGKDD*. ACM, 2013, pp. 989–997.

[29] R. W. White, M. Bilenko, and S. Cucerzan, "Studying the use of popular destinations to enhance web search interaction," in *ACM SIGIR*. ACM, 2007, pp. 159–166.

[30] R. Agrawal, R. Srikant *et al.*, "Fast algorithms for mining association rules," in *VLDB*, 1994, pp. 487–499.

[31] A. Mnih and R. Salakhutdinov, "Probabilistic matrix factorization," in *NIPS*, 2007, pp. 1257–1264.

[32] H. Ma, I. King, and M. R. Lyu, "Learning to recommend with social trust ensemble," in *ACM SIGIR*. ACM, 2009, pp. 203–210.

[33] M. Jamali and M. Ester, "A matrix factorization technique with trust propagation for recommendation in social networks," in *ACM RecSys*. ACM, 2010, pp. 135–142.

[34] S. Rendle, C. Freudenthaler, and L. Schmidt-Thieme, "Factorizing personalized markov chains for next-basket recommendation," in *WWW*. ACM, 2010, pp. 811–820.

[35] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, "Bpr: Bayesian personalized ranking from implicit feedback," in *UAI*. AUAI Press, 2009, pp. 452–461.