Interest-Based Product Recommendations along Temporal Dimension

Y. C. Liu*, P. Y. Hsu and S. W. Bai

Abstract — **Reed drastically, people can make connections and share information with friends easier and timely. Many researchers have noticed this subtle change and devoted to recommend products through WOM (word-of-mouth) on micro-blogs. However seldom of them consider that the popular topics and interests appeared in micro-blogs are changed frequently over time since the information is spread rapidly.**

In this paper, a new method to recommend products based on the changing preferences of individuals is proposed. Several experiments are done to verify the evolution of personal preference changes and to compare the recommendations made by top popular subjects, other domain experts and the personal preferences. The experimental results show the proposed method can recommend products more effectively.

Index Terms—**Product Recommendation, Micro-blog, Text Mining.**

I. INTRODUCTION

THE social network sites (SNSs) have boomed widely recently, SNSs nowadays attract lots visiting and become the hottest websites. Micro-blogs, such as: facebook, Twitter, Tumblr and Plurk, become more and more popular since they attract many participants to share information or interact with friends anytime anywhere.

Micro-blogs are commonly used to express personal opinions or feelings in time; people post messages to present their thoughts, intentions and opinions whenever they want. Consequently, it makes sense to collect and analyze the on-line posted messages to retrieve the product interests of bloggers; moreover the product recommendations can also be delivered on SNSs timely.

[1] pointed out the preferences and interests of consumers would be changed over time. Previous studies usually collect customer's demographic data as well as the product information to make product recommendations. However in the Information Era, customers can grab information at relative low cost and therefore change their interests frequently. The product recommendations made by previous work might be no longer applicable due to interests being changed frequently.

Existing product recommendation techniques highly reply on data collected from customers' purchasing histories, however in reality such information may not easily collected and does not provide the evolution of product preference changes of individuals.

As the micro-blogs nowadays are used to share personal feeling, intension and opinions timely; [1] has found that the messages posted on twitter can reflect the user's product preferences (interests) at those moments. Namely, the messages posted on micro-blogs provide valuable information to capture the changing product preferences of individuals timely.

Therefore, this work aims on developing approaches to detect the changing product preferences of customers first. By collecting on-line messages on micro-blogs, the evolutions of product preferences are captured. Further, the discovered, most update, product preferences are used to suggest customized product recommendations.

The proposed method first retrieve posted messages from micro-blogs, and then apply the text mining techniques to filter key words as customer's interests. Further besides the personal preferences, the expert opinions and the top most-update popular subjects are taken into considerations to form the product recommendations. The experimental results show the hybrid method can make better recommendations by comparing to previous methods.

The rest of this paper is organized as followed. In section II, related works are briefly reviewed. The proposed method to recommend products is detailed in Section III. Several experiments to testify the proposed method are shown in Section IV. Finally, the conclusion is drawn.

II. RELATED WORK

In this section, related literatures are briefly reviewed. Research on the effects of Word of Mouth in social network sites is discussed first. Secondly, literature of product recommendations on micro-blogs is introduced. Thirdly, the techniques commonly used in product recommendations are reviewed.

A. Word of Mouth in Social Network Sites

Word of mouth (WOM) is defined as passing information from person to person by oral communication, which is consisted of various forms such as: storytelling, oral tradition, oral history and word-of-mouth marketing…etc. The word-of-mouth marketing relies on the added credibility of person-to-person communication and is deemed as personal recommendation. [2][3]

Y. C. Liu is with the Department of Information Management, Shih Hsin University, Taipei, Taiwan 11604, ROC. (phone: 886-2-22368225-3362; fax: 886-2-2236-7114; e-mail: ycliu@mail.shu.edu.tw). (*corresponding author)

P. Y. Hsu is with the Department of Business Administration, National Central University, Jhongli, Taoyuan, Taiwan 32001, ROC. (e-mail: pyhsu@mgt.cc.edu.tw)

S. W. Bai is with the Department of Business Administration, National Central University, Jhongli, Taoyuan, Taiwan 32001, ROC. (e-mail: wunshow@yahoo.com.tw).

WOM marketing can be classified into several subcategories such as: buzz, blog, viral, grassroots, brand advocates, cause influencers and social media marketing. Through individual shares their satisfactions/dissatisfactions of certain products/services with other people, WOM is deemed as an effective marketing method. Because the communicators are not paid for their referrals, the individuals tend to believe the delivered information. [4]

[5] pointed that the key to successful WOM marketing is to locate influential endorsers to promote products; their work also concluded the on-line review comments written by so called "opinion leaders" are more effective than traditional advertisements. In general, the opinion leaders are with higher education background and social status so that they are more persuasive [6]. [4] further suggested to combine "opinion leaders" and WOM marketing together in promoting corporate products.

Since the social network sites nowadays are thought to provide an important platform for WOM marketing, [7] pointed that not only the opinion leaders but common users with lots friends can be very persuasive to recommend products. Therefore, the proposed method makes product recommendations without the limitation of opinion leaders; instead, opinions from friends within the same social groups are valued as influential recommenders.

B. Product Recommendations on Micro-blogs

Micro-blogs, such as twitter, facebook and Plurk … etc., are proved to make communications between friends easier and timely. As far, such social network sites have been one of the most popular sites which people visit every day. [8]

Nowadays, users get used to share facts and opinions which they interest in timely on SNSs. Since this on-line information comes from their known friends, people show more trusts on them. [9] further stated such trusts on on-line social networks can be used in marketing, and moreover corporate can grab customer's opinions and preferences timely. Hubspot also observed that more companies have invested more marketing expenses on micro-blogs/SNSs.

A well-known case is "DellOutlet". Dell is one of those companies that have gotten a lot of attention as an early adopter on twitter. They deliver promotion information of products to customers via twitter out of season and it has been proved to sale 3 million US dollars during 2007 to 2009 successfully. Excepting for products out of season, Dell also posts information and replies customer's questions via "DellOutlet" account on twitter frequently. Such on-line messages make two-way communication both effectively and efficiently.

Although "DellOutlet" has promoted products via micro-blogs well, it does not yet concern how to customize product recommendation based on individual's preferences.

C. Techniques on Recommender System

As the trend of big data analysis, the recommender systems are re-deemed as good instruments to filter out unnecessary information. Basically, the systems make recommendations according to customer's buying histories, preferences or behavior; the products are therefore categorized into customer groups to make recommendations [10]. Of no doubt, the recommender systems can be used in wide varieties.

In general, there are three main techniques used in recommender systems: collaborative filtering, content-based and the hybrid.

Collaborative filter mainly collects customer's purchasing histories and preferences, customers with similar purchasing experiences are grouped as clusters. According to purchasing preferences within the same group, common preferred products are therefore recommended to those customers not yet buying.

The content-based technique profiles customers by the products they bought or interested in. With considering the attributes or features of products which are bought (interested in) previously, similar products are recommended to the right customers. In brief, each customer is described by the attributes of products he/she bought previously, other products with similar attributes can be recommended to the corresponding customer.

The last but not the least, the hybrid method is proposed to combine the collaborative and content-based method, so that the drawbacks possessed by single method can be redeemed by the other. In brief, these techniques can be classified into two categories: the first one is merely to combine the recommendation lists made by both methods, and the other is to devise new algorithms combining both techniques together. Research has also proved that the hybrid method outperform previous work which applied single technique solely.

As stated above, although the collaborative filtering, content-based and hybrid method are the most common techniques adopted nowadays, there are still some problems needed to conquer. First, for general cases, the numbers of items in real world are huge; the products bought by customers are relatively small. The sparse interest matrix makes the similarities between users difficult to measure. Consequently, it's hard to decide which group customers belong and let alone recommend right products.

The "cold-start" is another noticeable problem. For a new customer, there is no information collected in databases. Hence, it's hard to decide which customer is his/her nearest neighbor so that no recommendation can be suggested.

In addition, the content-based technique summarizes the features of products which the customers bought to recommend similar products. However, it won't predict products which individual likes or is willing to buy. Moreover, products which can't be presented as "text content", such as: movie, music, travel … etc., are hard to be recommended through the content-based method.

In reality, the product preferences of individuals are frequently changed over time. [11] has already found the consumption preferences in retailing could be changed at different time points. Consequently, the micro-marketing strategy, the customized one-on-one product recommendations, is more suitable in the present business world. However, seldom work paid attention to such issues along temporal dimension. [12] pointed that the issues on detecting customer's preference changes are essential to on-line commerce companies especially since there are plenty valuable information provided on internet. As a result, to grab the current preferences of customers can help

companies target the right customers timely.

III. THE PROPOSED METHOD

In this section, the proposed method is introduced. There are three subsections. The flowchart of the proposed method is illustrated in Fig. 1. In the first subsection, the method to produce individual's (personalized) and expert's preferences is described. Later, the proposed approach for recommending products is detailed. Last, the way to capture the evolution of product preferences is explained.
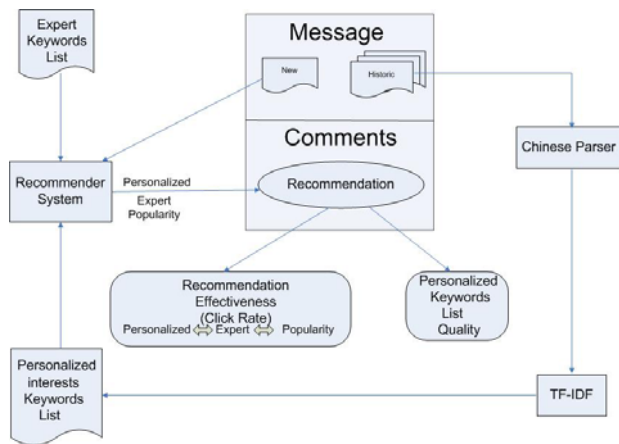

Fig. 1. The flowchart of the proposed recommender.

A. The Preference List: Individual's vs. Expert's

In this paper, the Chinese messages posted in micro-blogs are used to illustrate and experiment the proposed method, hence an additional step to parse Chinese terms is required. Of course, our method is not limited to Chinese documents; it can be applied to English documents freely.

Unlike English, there is no implicit space between Chinese words; hence it is necessary to break Chinese text into smaller linguistic units or segments, normally so-called Chinese terms [13]. In this study, the Chinese Parser developed by Science & Technology Policy Research and Information Center, National Applied Research Laboratories is adopted to parse the Chinese messages.

The proposed method first feeds messages (documents, in Chinese) into the Chinese Parser to retrieve the Chinese terms. Second, the well-known keyword retrieve technique is applied. The Term Frequency and Inverse Document Frequency (TF-IDF) is a widely adopted measure to screen keywords from parsed Chinese terms. [14] In this work, the keywords are proposed to represent the preferences/interests of users.

In brief, TF-IDF is a numerical statistic which measures the importance of a term to a collection of documents. It is used as a weighting factor in information retrieval and text mining. The TF value simply calculates the number of times a word appears in the document, but the IDF offsets the frequency of the word in all documents. The IDF value helps to filter out the words commonly appeared. The formulas are listed below.

$$TF(t,d) : f(t,d)/\max\{f(w,d), w \text{ in } d\}$$
$$IDF(t) : \log(|D|/|d \text{ in } D, t \text{ in } d|)$$
$$TF\text{-}IDF(t,d) = TF(t,d) * IDF(t)$$

In order to produce the individual's (personalized) preference list, the messages of certain user posted on micro-blogs are collected and fed into the Chinese parser. Further, a threshold k is set to screen top k keywords by descending TF-IDF values from the parsed Chinese terms. The top k keywords are therefore deemed as the personalized preference list.

Contrary to the personalized preferences, the preference list of the expert is provided by domain experts subjectively. Domain experts are supposed to have surveyed related messages broadly and therefore to identify important terms. Further, according to their knowledge, some terms which will rise sharply in near future are also considered into the preference list. This expert's list will further combine with the personalized list so as to make product recommendations more robust.

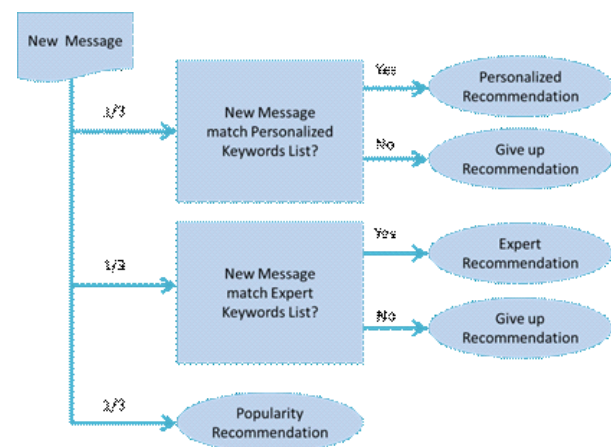#### B. The recommendation approach

Fig. 2. The proposed recommendation approach.

Fig. 2 depicted the main ideas of the proposed method. In brief, the proposed approach considers three various recommendations to form the final product suggestions. When a new message is posted on the micro-blog, the terms will be compared to the individual's or expert's preference list or the popularity trend randomly. Finally, the product recommendations are formed and posted as a reply to the original message. As a result, not only the user will see such recommendations but his/her friends will read them too.

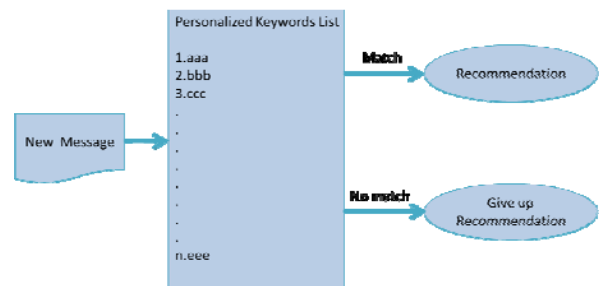Fig. 3 explains the product recommendations made by

Fig. 3. The personalized recommendation.

personalized keyword list described in previous subsection. The new posted message is supposed to be in Chinese, the message certainly needs to be parsed first. The parsed terms

are later compared to the individual preferences and the matched terms are suggested as product recommendations. The procedure to produce the expert's recommendations is almost the same and therefore omitted.

Moreover, the popularity recommender is built based on the top *N* clicked on-line advertisements. If the parsed terms from the new message match the keywords shown in top *N* advertisements, the corresponding product is recommended.

C. The evolution of preference changes

In order to capture the changing user's/expert's preferences, a mechanism to detect and update preference lists is required. The basic procedure is listed as Table 1.

Instead of updating the preference lists daily, the proposed method set a time window threshold *T* and a hit threshold *h* to keep the preference list fresh and timely. Basically, the preference list is periodical updated with time interval *T* illustrated as I-1 and II-1 in Table I. Meanwhile, the Time I-2(II-2) continues to collect the clicks of the proposed recommendations; if the count of clicks is less than the threshold *h*, the preference list would be re-calculated immediately.

TABLE I
THE TIME SCHEDULE FOR PREFERENCES UPDATING

Time	1			T			2T	1			3T
I-1											
I-2											
II-1											
II-2											

Fig. 4. The rank of visitor's nationalities on Plurk.

IV. EXPERIMENTS

In this section, the experiments are done to verify the effectiveness of the proposed method. The first subsection introduces the data used in experiments. Later, the evolution of the preference changing of individuals is verified. Last, the effectiveness of the personalized recommendations is shown.

A. The data preparation

The collection of customer's information

Since the Plurk is one of the most popular micro-blogs in

Taiwan, it is chosen to do the experiments. There are 35.6% bloggers on "Plurk" come from Taiwan, as shown in Fig. 4. Plurk is a free social networking and micro-blogging site in which users can send updates through short messages or links within 140 text characters in length. Users can respond to other users' updates from their timelines through the Plurk.com website, by instant messaging, or by text messaging. This study collects data from "Plurk" in Taiwan to identify the most update interests of bloggers to recommend suitable products timely.

The Plurk officially provides Python and Java API to develop recommender robots, several programs are coded in Java to collect the messages posted by each user. One hundred active users on Plurk are selected as the panelists.

The source of product recommendation

The products vended on the top one e-commerce site, books.com.tw, in Taiwan are adopted to do experiments for product recommending.

The number of selected keywords

Since the proposed method use the keywords to represent customer's preferences, the number of keywords should first be experimented. Fig. 5 shows the results, the number "37" is selected to do the experiments since it has the highest C/P value.
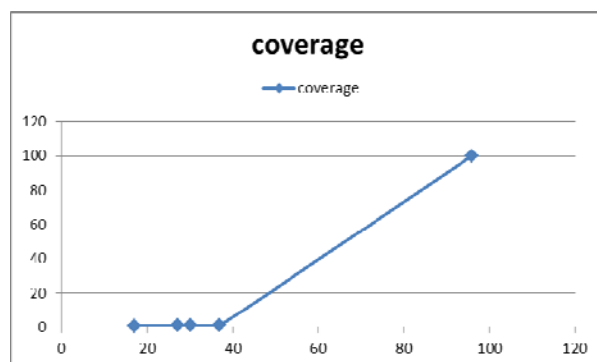

Fig. 5. No. of top keywords vs. coverage.

B. The observation of the changing preferences

In the experiments, the time windows are set to T=6 days. The click rate is defined as the equation listed below.

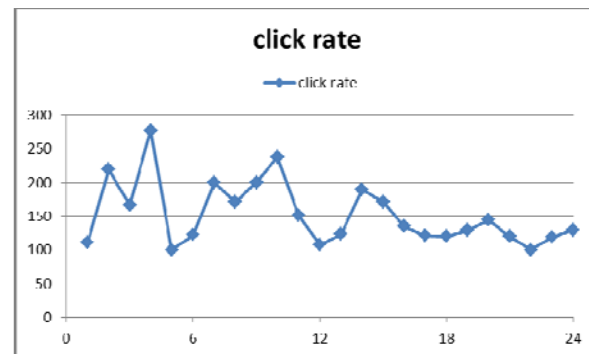Click rate = no. of clicks/no. of recommendations

Fig. 6. The click rates on four time periods.

The preference list is regularly updated every 6 days. From Fig. 6, it can be easily observed that the click rates increase soon after the newly updated preference is issued and decrease near the end of the current time window. Such observations further explain the user preferences will be

changed over time indeed.

C. The experimental results

Finally, the clicked rates of three recommendations are compared in Fig. 7. Instead of comparing the experimental results with other methods, we here merely compare the effectiveness of the personalized product recommendations with the expert and the popularity recommendations. It shows the personalized product recommendations outperform the other two methods.
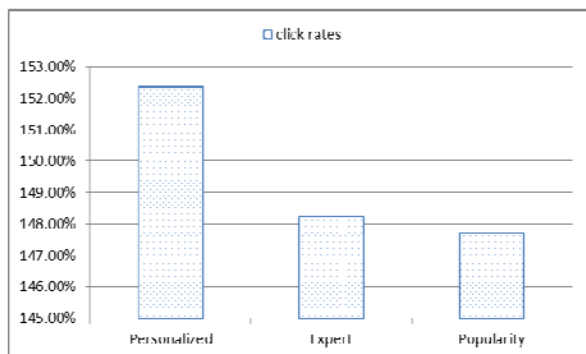

Fig. 7. The click rates comparisons.

V. CONCLUSIONS

In this paper, a hybrid method is proposed to use messages posted on micro-blogs as the preferences to make customized product recommendations. Unlike previous work, the frequently changed preferences are newly considered as an essential part of the proposed recommender system. Further, the experiments also show the messages posted in micro-blogs are testified to represent customer's product preferences well.

Although our work could make product recommendations effectively, there are still lots to do further. We are now working on more experiments to compare the proposed method with other work. Moreover, data collection from other social network platforms is also in progress.

REFERENCES

[1]   N. Banerjee, D. Chakraborty and K Dasgupta. (2009) User Interests in Social Media Sites: An Exploration with Microblogs.

[2]   E. W. Anderson. (1998). "Customer satisfaction and word of mouth," *Journal of Service Research. 1(1).* pp.5-17.

[3]   W. Mangold, F. Miller and G. Brockway. (1999). Word-of-mouth communication in service marketplace. *Journal of Service Marketing 13(1).* pp.73-88.

[4]   C. Timon and L. Feng. (2011). Who is talking? An ontology-based opinion leader identification framework for word-of-mouth marketing in online social blogs. *Decision Support Systems 51(1).* pp.190-197.

[5]   Yung-Ming Li, Chia-Hao Lin, Cheng-Yang Lai. (2010). Identifying influential reviewers for word-of-mouth marketing. *Electronic Commerce Research and Applications 9(4).* pp.294-304.

[6]   M. P. Venkatraman. (1989). Opinion leaders, adopters, and communicative adopters: a role analysis. *Psychology and Marketing 6(1).* pp.51-68.

[7]   T. Smith, J. R. Coyle, E. Lightfoot and A. Scott. (2007). Reconsidering Models of Influence: The Relationship between Consumer Social Networks and Word-of-Mouth Effectiveness. *Journal of Advertising Research 47(4).* pp.387-397

[8]   C.-J. Wang, S.-M. Cheng, L.-H. Lee, H.-H. Chen, W.-S. Liu, P.-W. Huang and S.-P. Lin (2012). NTUSocialRec: An Evaluation Dataset Constructed from Microblogs for Recommendation Applications in Social Networks. In Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC). pp. 2328-2333.

[9]   P.-S. Chang, I.-H. Ting, and, S.-L. Wang (2011). Towards Social Recommendation System Based on The Data From Microblogs. International Conference on Advances in Social Networks Analysis and Mining.

[10]   A.M. Rashid, I. Albert, D. Cosley, S. K. Lam, S. M. McNee, J. A. Konstan, and J. Riedl. (2002). Getting to know you : learning new user preferences in recommender systems, Proceedings of the IUT 02, San Francisco, CA, pp.127-134.

[11]   B. C. Yeong, , H. C. Yoon and H. K. Soung. (2005). Mining changes in customer buying behavior for collaborative recommendations. *Expert Systems with Applications 28.* pp.359-369.

[12]   S. H. Min and I. Han. (2005). Recommender systems using support vector machines, Proceeding of International Conference on Web Engineering, pp. 387-93.

[13]   S. Foo and H. Li. (2004) "Chinese word segmentation and its effect on information retrieval." *Information Processing and Management 40.* pp.161-190.

[14]   G. Salton and C. Buckley. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management 24(5).* pp. 513-523.