Webology, Volume 5, Number 3, September, 2008

Home	Table of Contents	Titles & Subject Index	Authors Index

A Personalized Word of Mouth Recommender Model

Chihli Hung
Department of Information Management, Chung Yuan Christian University, 200, Chung-Pei Rd., Chung-Li, Taiwan 32023 R.O.C.
Email: chihli (at) cycu.edu.tw

Received August 18, 2008; Accepted September 25, 2008

Abstract

Word of mouth (WOM) has a powerful effect on consumer behavior. Manually collecting WOM is very time-consuming in the era of the Internet. An automatic WOM recommender model is useful for both marketers and consumers. There are many different product features and thus many consumer choices. Each individual consumer has different preferences and these preferences may be changed deliberately or unwittingly. However, most existing WOM recommender models do not adapt to user preferences. This study proposes a conceptual WOM recommender model, which contains WOM collecting, document processing, recommending and user preference processing phases. More specifically, the self-organizing map (SOM) is used to store and abstract user preferences. This proposed WOM model makes recommendations to consumers or users according to their adaptive preferences.

Keywords

Consumer generated media; Buzz; Text mining; Sentiment analysis; Recommending agent; Self-organizing map

Introduction

Word of mouth (WOM) is oral person-to-person communication between an information receiver and a sender, who exchange the experiences of a brand, a product or a service based on a non-commercial purpose (Arndt, 1967). This means of communication serves as a crucial influencing role on consumer behaviors, such as buying, switching and diffusing (e.g. Arndt, 1967; Day, 1971; Sheth, 1971; Reingen, 1987; Brown & Peter, 1987; Murray, 1991). From the marketer viewpoint, how to harness and manage opinions from customers is very important. However, the purpose of WOM is not for commerce or marketing. Product information provided by a manufacturer is usually seen as a marketing approach, which has less effect on their consumers than information provided on a consumer opinion website or a discussion board on the Internet (Bickart & Schindler, 2001).

In the era of information, the Internet provides human beings with a new way of communication, which is more efficient and effective. Word of mouth may be in electronic form, diffused via the Internet. An information sender is able to communicate not only with family, friends and acquaintances, but also with anyone who may be unknown to the sender but has some interest in the information. This is true also for an information receiver, who is able to obtain information from anyone on the Internet. Thus, word of mouth on the Internet influences the information receivers more quickly, broadly, widely, significantly and without any geographic limitation (Goldsmith, 2006).

Nowadays, a potential buyer may take advantage of the power of WOM before making a decision (Arndt, 1967; Leskovec et al. , 2007). By referring to online articles, a buyer is able to save decision-making time and may make a better buying decision (Hennig-Thurau, 2004). However, the mass of information available on the Internet has caused users to encounter information chaos (Davenport & Beck, 2002). How to generate the relevant information has become an essential topic. Some word of mouth mining tasks only focus on a single resource, such as a single web site (e.g. Hu et al., 2007; Abbasi et al. , 2008), which seems to be less useful in a real world environment. Others extract word of mouth textual information through multiple Internet resources (e.g. Endo & Noto, 2003).

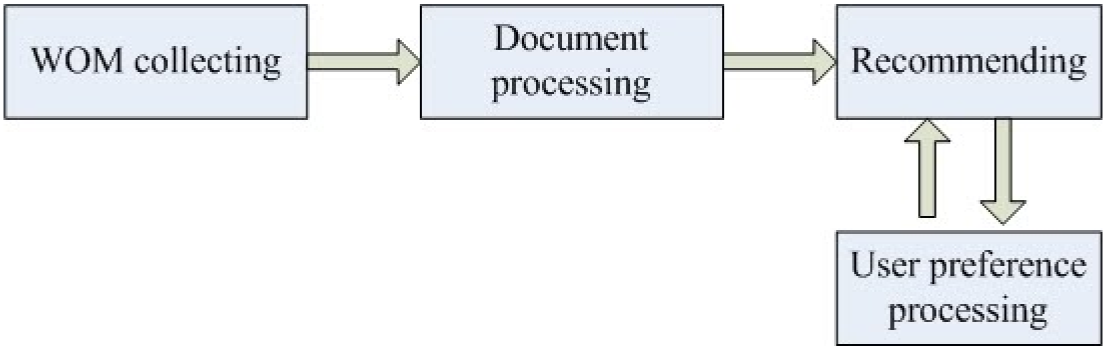
For word of mouth mining, a specific product or service has its specific features. On the other hand, a specific potential buyer has his/her own specific buying preferences. For example, a specific product, such as a mobile phone, contains several product features, such as price, handiness, battery life, durability, and so forth. The most important product feature for potential buyer A is price, while for potential buyer B it is handiness. That is, the order of the significant buying features for different buyers may be different. Cheung et al. (2003) mined customer product ratings for personalized marketing. Endo & Noto (2003) built a WOM information recommender system by considering information reliability and user preferences. However, they did not consider the sentiment analysis and the sentiment measure for WOM. Most existing word of mouth mining models do not consider the different preferences of their users (e.g. Straub & Heinemann, 2004; Aciar et al., 2006). In particular, they do not learn from accumulating the experiences within users' preferences.

This paper proposes a conceptual model which mines personalized word of mouth textual information. This proposed model stores users' dynamic experiences in the self-organizing maps (SOMs) (Kohonen, 1982) as users' interest bases, which compare and re-rank the word of mouth mining results.

The Proposed Personalized Word of Mouth Recommender Model

The proposed personalized WOM recommender model consists of four phases, which are WOM collecting, document processing, recommending and user preference processing (Figure 1). In the WOM collecting phase, the proposed model collects relevant WOM information for a given product. The main issues for this phase are how to locate the relevant WOM information and how to filter out irrelevant information. In the document processing phase, the proposed model firstly acts as a traditional text mining model for WOM. The proposed model then analyzes the sentiment concept for WOM documents. In the recommending phase, the proposed model ranks the relevant products based on users' preferences. However, this proposed model becomes a normal WOM recommender model when there is no users' preference information at the very beginning. In the user preference processing phase, the proposed model feeds the user post-search processing into two preference bases, which are formed by a self-organizing map (SOM). The recommended product list is then re-ranked based on these preference bases.

Figure 1. A conceptual architecture of the personalized WOM recommender model

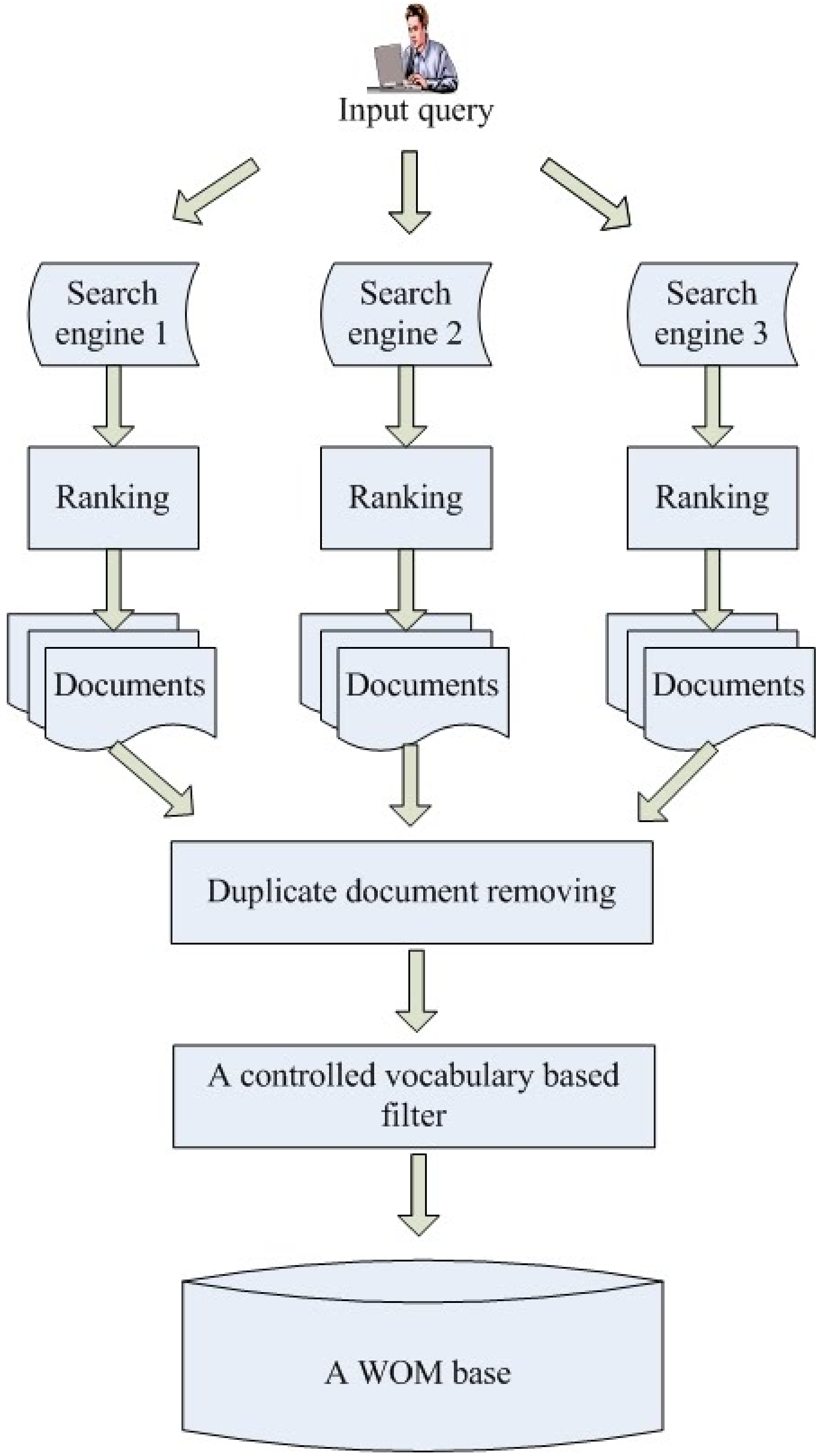


Word of Mouth Collecting

The initial requirement of a potential buyer is to search for WOM for the relevant products. How to collect useful WOM is the first main issue for the design of a personalized WOM recommender model. WOM can be stored in various web sites and in different electronic formats (Wang et al., 2005). For example, some WOM information exists on sites which are dedicated to a specific type of product such as mobile phones, travel agents, and so forth. Some WOM information exists on the sites of newspapers and magazines which may contain product reviews. Some WOM information may exist directly on the discussion board of an electronic shopping mall. Some WOM information may exist on sites that are designed specifically to collect user reviews for specific types of product. Some WOM information also exists on personal web blogs. In the age of the Internet, search engine technique has a dominant influence for searching relevant information (Berry & Browne, 2005). A meta-search strategy, which produces a higher recall rate than a single search engine (Lehikoinen et al., 2006), is also used in this research. In particular, Google, Yahoo and MSN search engines are used for the same query. The conceptual illustration for this phase is shown in Figure 2.

For each search engine, we extract a pre-defined number of web pages, for example 500 web pages in each ranking list. As each search engine has its own ranking algorithm, the meta-search strategy is able to collect a greater volume of relevant information. Some information may be duplicated by a meta-search engine. We only reserve one copy of such information. However, when trying to locate WOM on a given product, a meta-search engine usually turns up several relevant web pages which are mixed with many irrelevant ones. For example, some pages do not contain the name of the given product or do not evaluate the given product. Dave et al. (2003) crawled search engine results and removed any web pages without the word "review" in their title. We also filter non-WOM information from the results of the meta-search engine. Based on the appraisal theory (Martin & White, 2005), which clearly defines the sentiment properties of language that can be used for WOM mining, we build a controlled vocabulary base so as to evaluate WOM sites for the specific product or service.

Figure 2. The word of mouth collecting phase of the proposed model



Word of Mouth Document Processing

The main purpose of this phase is to give each linguistic unit of WOM a sentiment tag. A linguistic unit (LU) consists of four parts in our definition. The first part of the LU is a product feature. The second part of the LU is a modifier for this product feature. The third part of the LU is a sentiment orientation and the fourth part of the LU is a measure of the sentiment description. The conceptual diagram is shown in Figure 3.

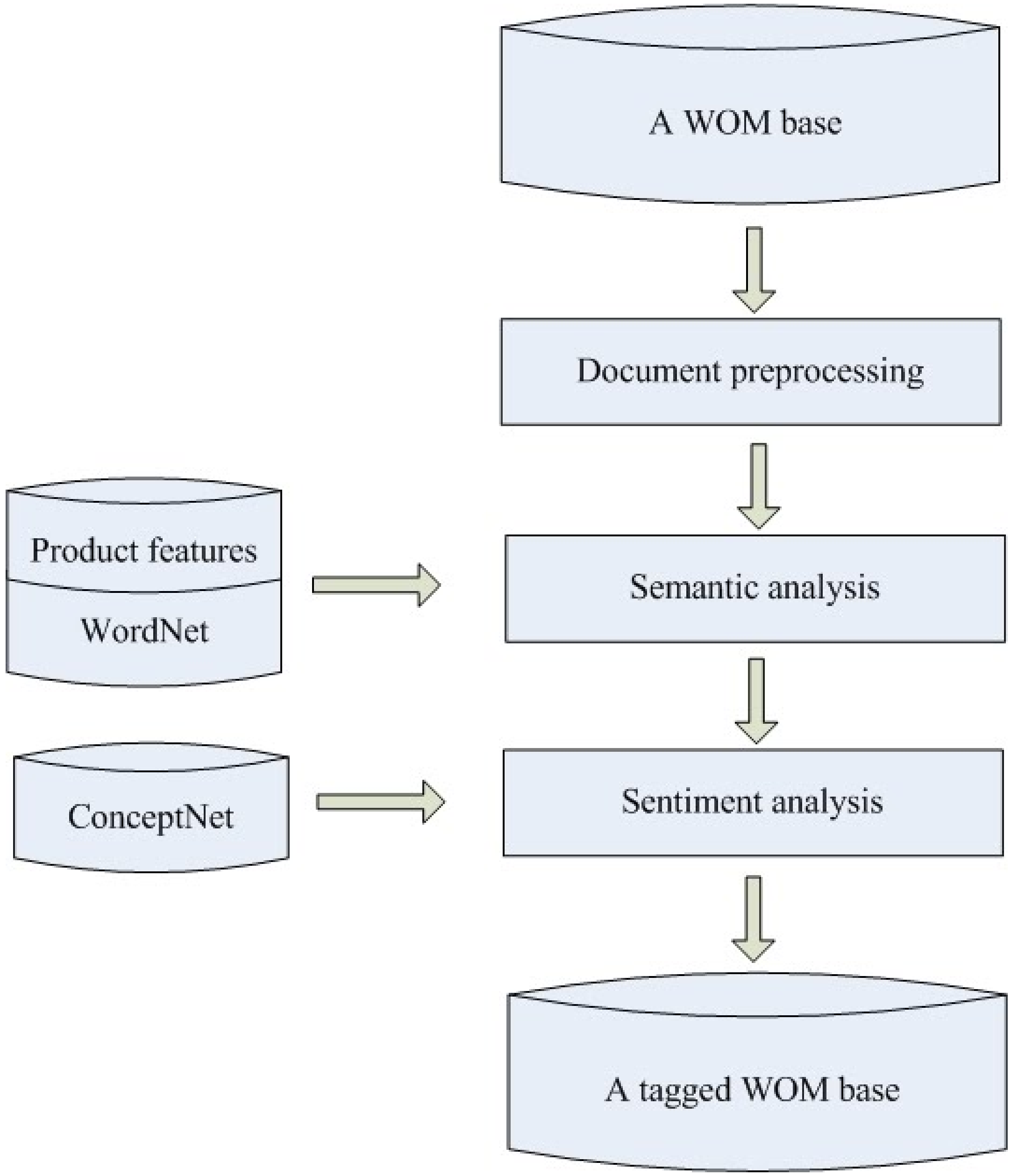
As the core part of the LU is a product feature, we need to extract product features from a description of the product. However, this description on the Web is usually provided by a manufacturer, which may not satisfy the needs of a user. One possible way to collect significant product features is via user questionnaires. The product features can also be collected by going through the marketing literature on the given product. For instance, Mazzoni et al. (2007) suggested that features of mobile phones consist of four categories, which are economic factor, physical factor, aesthetical factor and technological factor. Each factor contains some detailed features. For example, "price", "promotion" and "tariff" belong to the economic category. The physical category contains "handiness", "battery life", "screen visibility", "durability" and "signal reception". The aesthetical category contains "aesthetics", "personalization" and "brand reputation". The technological category contains "advanced services", "accessories" and "other functions". In a questionnaire more concrete examples are given to describe obscure features. For example, the personalization feature includes "cover", "ring tone", and so forth.

We extend the original product feature term by looking up the WordNet ontology (Miller, 1985). All synonyms in one level of the hypernym tree and the hyponym tree are considered relevant product features. Thus we compose an extended product feature list. In order to choose a proper sense for a word in the WordNet, we compare the gloss in each sense for a given word with the description of the feature in the questionnaire.

On the other hand, we remove all html and xml tags from WOM web pages in the WOM base and use the part of speech tagger (Brill, 1992) for the extraction of subjective nouns, objective nouns and adjectives. We then fill a noun in the field of the product feature of the LU when such a noun is found in the extended product feature list. However, when this noun is a pronoun, we search previous sentences until a concrete noun is found. We next fill the adjective located after a subjective noun in the LU and the adjective located before an objective noun. When a subjective or objective noun is associated with more than one adjective, the previous feature in the LU is used.

In order to fill in the sentiment orientation of LU, we use the ConceptNet API function, such as the GuestMood function, which is able to classify a concept into one of six affect categories, such as happy, sad, angry, fearful, disgusted and surprised (Liu & Singh, 2004). When a negative modifier, such as "not", is used, the sentiment orientation turns to the opposite. However, if a negative prefixed modifier, such as "un", or negative postfix, such as "less", is used, the ConceptNet handles it directly. As the ConceptNet provides an incomplete coverage of words, we make use of the WordNet ontology by looking up the synonym set level by level in order to extend its scale. For filling in the fourth part of the LU, such as a measure of the sentiment description, we analyze the usage of word and give each special usage greater weight. For example, a comparative form of adjective has more weight than its original form and a supreme form of an adjective has more weight than its comparative one. A concept founded in a title sentence should have greater weight. A concept founded in a conclusion sentence should also have greater weight.

Figure 3. The word of mouth document processing phase of the proposed model



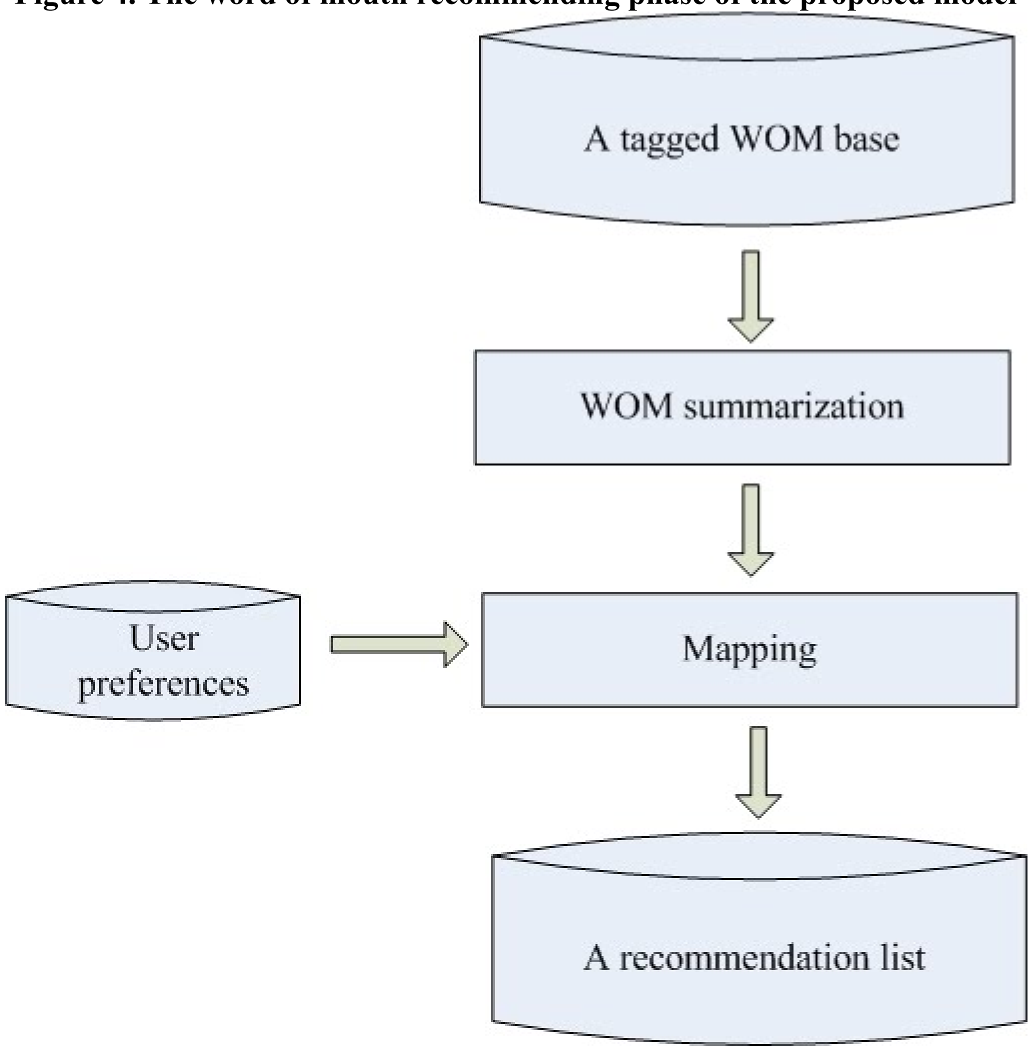
Word of Mouth Recommending

After the previous phases, we have a tagged WOM base, which contains many linguistic units. Manually collecting overall sense of those linguistic units is still time-consuming, so we apply the technique of document summarization. Unlike traditional document summarization, which expresses the concept of the original long text by a shorter document, WOM summarization focuses on the overall sentiment of a large amount of WOM (Lee et al. , 2008). As the tagged WOM base is constructed by many linguistic units, which contain product features, modifiers, sentiment orientation and measures of sentiment, to achieve an overall sentiment is straightforward. The main issue in this WOM recommending phase is to rank the relevant products based on user preferences. We show our conceptual illustration in Figure 4.

User preferences are collected by observing how the user evaluates the recommendation list provided by our model. We divide the user preferences into two sub-preferences according to sentiment orientation. That is, we have one positive user preference base and one negative user preference base. Both user preference bases are built by utilizing the technique of the self-organizing map (SOM), which is regarded as the user endogenous attention (Hung et al., 2008). The SOM is able to represent multi-dimensional data onto a two-dimensional output map for visualization, which provides a prominent abstraction ability (Kohonen, 2001). We use this abstraction feature to extract the most significant user preferences.

The positive SOM expresses what the user likes while the negative SOM expresses what the user dislikes. By mapping the tagged WOM base with these two SOMs respectively, a recommendation list for a given product can be provided. More specifically, we treat the positive SOM as a user query for the tagged WOM base. On the other hand, the negative SOM is considered a user constraint to filter out the search results.

Figure 4. The word of mouth recommending phase of the proposed model



User Preference Processing

The main aim of the user preference processing phase is to build and maintain two preference bases for individuals, such as a positive preference base and a negative preference base. By observing the post-behavior on the recommendation list, language units in a tagged base are collected. The conceptual diagram is shown in Figure 5. At the user evaluation stage, the proposed model provides a recommendation list based on user positive preferences, and a non-recommendation list based on user negative preferences. The user picks up some recommended and non-recommended products via a user interface. The model then searches associated language units in the tagged WOM base for such relevant products.

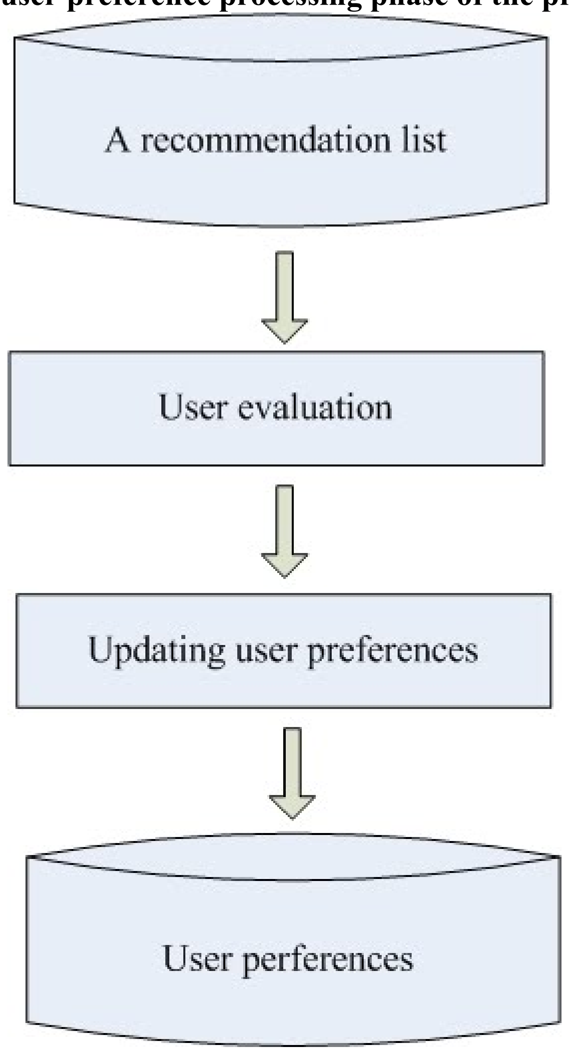
For each chosen product, we combine the product feature and its modifier as a combined attribute, and its measure of the sentiment description in language units is the value of this combined attribute. We simply fill in the attribute and its measure in Table 1. This table is then trained by the SOM algorithm.

As the SOM is truly an unsupervised learning approach, it is necessary to assign meaningful labels to units. An attribute term usually represents input samples because this term is important for those samples. Based on the vector space model in the field of text mining, a more important term is usually transformed to a greater element value (Salton & Buckley, 1988). Therefore, the element with a greater value in a unit vector can be used as a representative term for input samples, which are mapped to this unit (Ritter & Kohonen, 1989; Roussinov & Chen, 1999).

Table 1. An example of user preference matrix

	Attribute 1	**Attribute 2**	**. . .**	**Attribute n**
Sample 1	3	0	. . .	0
Sample 2	3	2	. . .	4
. . .	. . .	. . .	. . .	. . .
Sample m	0	4	. . .	1

Figure 5. The user preference processing phase of the proposed model



Conclusions

WOM has been proven to be of great importance for both marketers and consumers (Arndt, 1967). This paper proposes a conceptual architecture of a personalized WOM recommender model, which contains four phases, namely WOM collecting, document processing, recommending and user preference processing. The WOM collecting phase collects relevant WOM information via a meta-search engine. The document processing phase analyzes the sentiment concept for WOM documents. The recommending phase ranks the relevant products based on users' preferences. The user preference processing phase builds and maintains user preference bases to adapt to user intentions, which are formed by a self-organizing map. The main novel aspect of this paper is the method for storing, maintaining and extracting the user preferences in self-organizing maps adaptively. Therefore, this proposed WOM model recommends to users according to their adaptive preferences.

References

- Abbasi, A., Chen, H., & Salem, A. (2008). Sentiment analysis in multiple languages: feature selection for opinion classification in web forums. *ACM Transactions on Information Systems*, 26(3).
- Aciar, S., Zhang, D., Simoff, S., & Debenham, J. (2006). Informed recommender agent: utilizing consumer product reviews through text mining. *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*.
- Arndt, J. (1967). Role of product-related conversations in the diffusion of a new product. *Journal of Marketing Research*, 4(3), 291-295.
- Berry, M.W., & Browne, M. (2005). *Understanding Search Engines*, SIAM, PA, USA.
- Bickart, B., & Schindler, R.M. (2001). Internet forums as influential sources of consumer information. *Journal of Interactive Marketing*, 15(3), 31-40.
- Brill, E. (1992). A simple rule-based part of speech tagger. *Proceedings of the Third Conference on Applied Natural Language Processing*.
- Brown, J.J., & Peter, H.R. (1987). Social ties and word-of-mouth referral behavior. *Journal of Consumer Research*, 14(3), 350-362.
- Cheung, K.-W., Kwok, J.T., Law, M.H., & Tsui, K.-C. (2003). Mining customer product ratings for personalized marketing. *Decision Support Systems*, 35, 231-243.
- Dave, K., Lawrence, S., & Pennock, D.M. (2003). Mining the peanut gallery: opinion extraction and semantic classification of product reviews. *Proceedings of the 12th international conference on World Wide Web*, Budapest, Hungary, pp. 519-528.
- Davenport, T.H., & Beck, J.C. (2002). *The Attention Economy: Understanding the New Currency of Business*, Harvard Business School Press, September 2002.
- Day, G.S. (1971). Attitude change, media and word of mouth. *Journal of Advertising Research*, 11(6), 31-40.
- Endo, H., & Noto, M. (2003). A word-of-mouth information recommender system considering information reliability and user preferences. *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, pp. 2990-2995.
- Goldsmith, R.E. (2006). Electronic word-of-mouth. In Mehdi, K.-P. (Ed.), *Encyclopedia of E-Commerce, E-Government and Mobile Commerce* (pp. 408-412). Hershey, PA: Idea Group Publishing.
- Hennig-Thurau, T., Gwinner, K.P., Walsh, G., & Gremler, D.D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *Journal of Interactive Marketing*, 18(1), 39-52.
- Hu, N., Liu, L., & Zhang, J. (2007). Analyst forecast revision and market sales discovery of online word of mouth. *Proceedings of the 40th Annual Hawaii International Conference on System Sciences* (*HICSS'07*).

- Hung, C., Chi, Y.-L., & Chen, T.-Y. (2008). An attentive self-organizing neural model for text mining. *Expert Systems with Applications*, doi: 10.1016/j.eswa.2008.08.037.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, 59-69.
- Kohonen, T. (2001). *Self-organizing maps*, 3rd edition. Springer-Verlag, Berline, Heidelberg, New York.
- Lee, D., Jeong, O.-R., & Lee, S.-G. (2008). Opinion mining of customer feedback data on the Web. *Proceedings of the 2nd international conference on Ubiquitous information management and communication*.
- Lehikoinen, J., Salminen, I., Aaltonen, A., Huuskonen, P., & Kaario, J. (2006). Meta-searches in peer-to-peer networks. *Personal and Ubiquitous Computing*, 10(6), 357-367.
- Leskovec, J., Adamic, L.A., & Huberman, B.A. (2007). The dynamics of viral marketing. *ACM Transactions on the Web*, 1(1).
- Liu, H., & Singh, P. (2004). ConceptNet: a practical commonsense reasoning tool-kit. *BT Technology Journal*, 22(4), 211-226.
- Martin, J.R., & White, P.R.R. (2005). *The Language of Evaluation: Appraisal in English*. Palgrave, London.
- Mazzoni, C., Castaldi, L., & Addeo, F. (2007). Consumer behavior in the Italian Mobile telecommunication market. *Telecommunications Policy*, 31, 632-647.
- Miller, G.A. (1985). WordNet: a dictionary browser. *Proceedings of the First International Conference on Information in Data*, University of Waterloo, Waterloo, pp. 25-28.
- Murray, K.B. (1991). A test of services marketing theory: consumer information acquisition activities. *Journal of Marketing*, 55(1), 10-25.
- Reingen, P.H. (1987). A word of mouth network. *Advances in Consumer Research*, 14(1), 213-217.
- Ritter, H., & Kohonen, T. (1989). Self-organizing semantic maps. Biological. *Cybernetics*, 61, 241-254.
- Roussinov, D.G., & Chen, H. (1999). Document clustering for electronic meetings: an experimental comparison of two techniques. *Decision Support Systems*, 27, 67-79.
- Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5), 513-523.
- Sheth, J.N. (1971). Word-of-mouth in low-risk innovations. *Journal Advertising Research*, 11(3), 15-18.
- Straub, T., & Heinemann, A. (2004). An anonymous bonus point system for mobile commerce based on word-of-mouth recommendation. *Proceedings of the 2004 ACM Symposium on Applied Computing*, pp. 766-773.
- Wang, J.-C., Chiu, C.-C., & Tang, J.-J. (2005). The correlation study of eWOM and product sales predications through SNA perspectives: an exploratory investigation by Taiwan's cellular phone market. *Proceedings of the 7th International Conference on Electronic Commerce* (*ICEC'05*), Xi'an, China.

Bibliographic information of this paper for citing:

Hung, Chihli (2008). "A personalized word of mouth recommender model." *Webology*, **5**(3), Article 61. Available at: http://www.webology.org/2008/v5n3/a61.html

---
