Identifying Bloggers with Marketing Influence in the Blogosphere

Yung-Ming Li
Institute of Information Management, National Chiao Tung University
1001 University Road, Hsinchu, Taiwan 300, ROC
886-3-5712121 Ext. 57414
yml@mail.nctu.edu.tw

Cheng-Yang Lai
Institute of Information Management, National Chiao Tung University
1001 University Road, Hsinchu, Taiwan 300, ROC
886-3-5712121 Ext. 57427
cylai.iim96g@g2.nctu.edu.tw

Ching-Wen Chen
Institute of Information Management, National Chiao Tung University
1001 University Road, Hsinchu, Taiwan 300, ROC
886-3-5712121 Ext. 57427
kisd0955@hotmail.com

ABSTRACT
Finding influential bloggers will not only allow us to better understand interesting activities happening in a social network, but also present unique opportunities for sales and advertisements. In this paper, we address a novel problem of finding influential bloggers with marketing value in the blogosphere by proposing a MIV (Marketing Influential Value) model. We induce two dimensions of blog characteristics (network-based factors and content-based factors) and develop an adaptive artificial neural network (ANN) to identify the potential bloggers with marketing influence to support the marketers or advertisers in promoting their products or services effectively.

Categories and Subject Descriptors
I.5.2 **PATTERN RECOGNITION**: Design Methodology – *Classifier design and evaluation, Feature evaluation and selection, Pattern analysis*.

General Terms
Measurement, Design, Experimentation, Human Factors.

Keywords
Influential model, viral marketing, social networks, blogosphere

1. INTRODUCTION
With the advent of online social network, word-of-mouth (or viral) marketing is increasingly being recognized as a crucial strategy in social influence and marketing domains. Through word-of-mouth diffusion, information could spread more quickly and easily among social networks. The essence of word-of-mouth marketing is to reach out to a broad set of potential customers and attract considerable attention via social interactions. Unlike direct and mass marketing which only recognize the intrinsic value of a customer, word-of-mouth marketing additionally exploits the network effect of customer by taking the network value into consideration to measure the real customer value [24]. Appropriate marketing campaigns developed based on the blog social network could generate significant performance in increasing the sales and reducing the promotion costs.

Blogging systems has gained much attention as a merging social medium, which exploits existing social networks by inspiring bloggers to share their own posts or personal information with other Internet surfers. The weblogs indeed provides a more open channel of communication for people in the blogosphere to read, commentate, cite, socialize and even reach out beyond their social networks, make new connections, and form communities [19]. Blogging is one huge word-of-mouth engine [25] and blogosphere has also become a finest platform for the advertisers to promote a new product or service and the customers to locate product comments and purchasing suggestions. In this research, we are particularly intrigued by the issues in modeling the key factors which contribute to a successful marketing influencer and identifying the potential marketing influencers in the blogosphere.

A typical blog site combines text (basic content), images or video (multimedia) and a variety of links (network-based linkage). In our work, these factors are categorized into two dimensions – network-based and content-based factors. Both network-based and content-based sources should be considered to develop a more comprehensive and robust influence model and estimate precise value of marketing influence. In this research, social influence is modeled as a graph-based representation, which is formed by nodes and edges. Nodes stand for blog sites on the blogosphere with some social characteristics or behaviors and have social influence value. Edges represent directed social influence. Considering nonlinear and complexity characteristics of blogging behaviors, we utilize the artificial neural network (ANN) approach to predict the overall influential power of the targeted bloggers. The proposed model was based on Wretch - a prestigious online blogging system with largest number of bloggers in Taiwan.

The rest of this paper is organized as follows. Section 2 reviews related works to the study. Section 3 investigates the factors of marketing influence to develop the MIV model. Section 4 elaborates on our MIV model formed by network-based value and content-based value. Section 5 proposes an experimental study to empirically understand the MIV model. Section 6 concludes our findings in this research.

2. RELATED WORKS
2.1 Blog Social Network: Ranking mechanism
A fast-growing number of blog studies have shown that blogosphere as a social network can help researchers in

ICEC '09, August 12-15, 2009, Taipei, Taiwan

understanding and analyzing certain implications and insights. Prior research demonstrated social relation-based dimension to measure the importance and relationships of webpages or blogs [2, 10, 20, 23]. The concept of blog ranking is similar to that of blog recommendation to some extent. Fujimura et al. [10] assigns scores to each blog entry by weighting the hub and authority scores of the bloggers based on eigenvector calculations, which has similarities to PageRank [4] and HITS [18] in that all these studies are based on eigenvector calculation of the adjacency matrix of the links. However, the work in Kritikopoulos et al. [20] ranks blogs according to their similarity in social behaviors by graph-based link analysis, which demonstrates an excellent paradigm of link analysis. Note that there is an inherent problem of sparseness in the blogosphere which has already been noticed by researchers. Adar et al. [2] and Kritikopoulos et al. [20] have coped with this problem by extending and increasing explicit and implicit links based on various blog aspects where a denser graph will result in a better performance of ranking and recommending.

2.2 Social Influence of Blog

Finding influential blog sites in the blogosphere is an important research problem, which investigates how these blog sites influence the external world and within the blogosphere [11]. The works [3, 15, 26] applied various network-based parameters and dimensions to examine the influence of commentary information in online social network. Java et al. [15] applies simpler, PageRank-based heuristics for influence models on graph derived from links between blogs and between general web pages. They also discuss the applicability of the proposed algorithms; this gives us a great demonstration on how to justify and verify a proposed heuristic or algorithm. Moreover, Agarwal et al. [3] presents a well-defined preliminary influence model to identify influential bloggers and paves the way for building a robust model that allows for finding various types of the influentials. In our model, we extend the concepts and ideas from Agarwal et al.[3], however with a different problem definition in the online marketing domain. Identifying the blog sites with greater marketing influence capabilities is crucial in promotion of a product/service.

2.3 Blog Marketing via Word-of-mouth

In a social network, marketing through word-of-mouth effect is extremely powerful. People are likely to be affected by the decisions of their friends and colleagues [16], and some researches have investigated the diffusion process of word-of-mouth and viral marketing effects in the success of new products [8, 24]. Zhan et al. [31] emphasize the important role of writing and referring product reviews in the internet (such is blogosphere or online communities). In the case of the methodologies to implement opinion mining, many scholars focus on the identification of author's attitude such as positive or negative [8]. Motivated by applications to marketing, several probabilistic models are proposed for choosing customers with a large overall effect on the social network [8, 24]. Notably, the problem is addressed that we hope to market a new product/service via the adoption of the power of word-of-mouth in the network. Or say we start by initially targeting a few influential members as the virus of the network to diffuse and propagate the information even recommendations to their friends [16]. But how should we choose the influential seeds with the strongest virulence? Or which characteristics should be taken into consideration for indentifying the influential nodes?

Our work is based on the problem motivated by Agarwal et al [3]. and Kempe et al. [16] and influence models proposed by Agarwal et al. [3], Subramani et al. [26], and Java et al. [15]. These models aim to indentify and model the spread of influence in online social networks. However, this research focuses on identifying the nodes with marketing influence in the blogosphere and differs from those briefly discussed above. In the next section, we examine several factors of marketing influence to construct our MIV (marketing influence value) model.

2.4 Back-propagation Neural Network

Artificial neural network (ANN) is a model which is utilize mathematical or computational model based on biological neural networks. The researches have already been show that the learning ability of back-propagation neural network to conduct forecast and prediction is appropriate in different domains [5, 21]. The major advantage of neural network is the flexible nonlinear modeling capability. The purpose of ANN is to construct a model which could learn weights similar to the human thinking. Kuo and Chen [21] applied fuzzy neural network to learn the rules produced from order selection questionnaires in electronic commerce. A feed-forward ANN with error back-propagation learning algorithm is also employed to integrate different scores. Chiang et al. [5] developed an ANN model to predict and explain consumer's choice between web and physical stores.

3. MARKETING INFLUENCE VALUE (MIV) Model

In this section, we propose a marketing influence model to determine a set of potential bloggers with high market influence. The role of the blogger is mainly to create awareness and signal benefits to others within their blog social network; they can be particularly influential in encouraging trial and adoption of novel products and services [26]. In this model, we divide the marketing influence value into two main categories, which are network-based and content-based values, as shown in Fig 1.

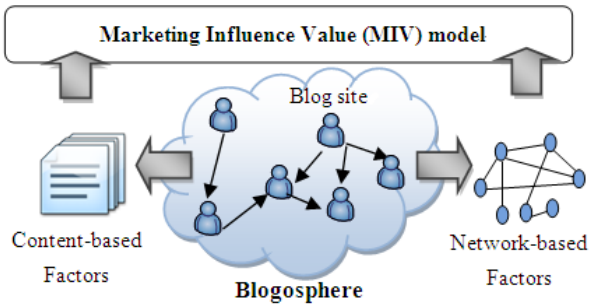

Figure 1. Social influence and MIV factors on the blogosphere

3.1 Network-based Factors

Blogosphere incorporates several important factors and properties in examining whether a blogger indeed have sufficient potential to elaborate the influence of viral marketing. These important factors are described as follows.

- *In-links and out-links*

The nature of blog social network forms by linkage structure, including in-degree and out-degree links. Although there is an inherent problem of sparseness in the blogosphere, Adar et al. [2] and Kritikopoulos et al. [20] cope with it by extending and increasing explicit and implicit links based on various blog aspects, where a denser graph will result in a better performance of ranking and recommending. According to Agarwal et al. [3], the number of out-links denotes negative utility in influential value and loss in novelty (originality) of a blog post. Brin and Page [4] also suggested that the PageRank score contributing to other nodes should be divided by the number of out-links. In-links structure in blog marketing domain suggests the concept of brand and recognition, which means the more in-links shows stronger brand power and well recognized blog content. The PageRank algorithm works on a web graph by simulating a random walk with a probability q to jump to other sections of the web graph and probability $1-q$ to follow a link and. Denote $IL(a)$ as the set of in-links of web page *a* and $OL(a)$ as the set of out-links of web page *a*. PageRank score of web page a, $PR(a)$, is measured as:

$$PR(a) = q + (1-q)\sum_{p_i \in IN(a)} PR(p_i) / |OL(p_i)|,$$

where $|OL(p_i)|$ are the number of out-links of page p_i. Based on this concept of in-links and out-links, we integrate the concept of social network. The in-links and out-link value respectively denotes the number of friend-of and the number of friends in the social network for one blogger.

- *Comments and citations*

To measure the impacts on viral marketing effect, we take the notions used in Agarwal et al. [3] and deem the number of comments it received as the capabilities of generating activities. In other words, few or no comments indicate little interest of fellow bloggers to heed his/her advice of purchasing a particular product. Citation is a kind of in-degree link which could be taken as a rough substitute for recognition level of a blog post. Thus, the marketing influence in adopting products/services or propagating epidemic information should be more significant when a blogger/blog post have more citations then others.

- *Network externalities*

Network externality is a positive utility. In addition to those directly derived from social network usage, network externality accrues broadly to the set of all adopters [26]. Therefore, the number of visitors (α) during a period of time in a blog site was taken as an indicator for the network externality factor. These implicit network links are created by the visiting behaviors. It's quite intuitive that a blog site can be more effective in promoting product/service when it has more disseminators.

- *Blogroll relationship (Trustworthiness)*

Friend or friend-of relationship is especially a crucial factor in referencing the trustworthy and reliable information [22]. In blog marketing domain, the "reputation" (γ) was taken to measure the trustworthiness of a blogger. Various definitions and notions of trust and reputation have been studied across diverse disciplines. Previous studies [1, 27] indicated that the reputation could been seemed as a form of social control mechanism due to users will fear of gaining bad reputation. In our context, the reputation factor is defined as a similar meaning, a perception that an agent creates through past actions about its intentions and norms.

- *NV (Network based Value) calculation*

Based on the above equation, we transform the web context into the blog social network - a representation of marketing influence graph in which the number of friend-of (IL) and the number of friend (OL) are taken into consideration for calculating network-based value. Network-based value ($NV(b)$) of blog site b is affected by the above factors where number of visitors (α), reputation (γ), number of comments(c), and number of citation (η), are taken into the equation to prepare a ground for developing MIV model. Thus, we combine these factors to determine the network-based value of blog site b,

$$NV(b) = \alpha_b(PR(b)) + \gamma_b f(c_b, \eta_b),$$

. Note that, the popularity effect implies the degree of recognition, trustworthiness and hotness of a blog site and $f(c_b, \eta_b)$ is a function of popularity effect from c_b and η_b.

3.2 Content-based Factors

Since blog is formed by text-based articles with reverse chronological sequences of dated entries, a content-oriented analysis is indispensable to better understand the tendency, preference and viewpoint of bloggers about the marketing/advertising of specific product/service. The important content-based factors are described as follows.

- *Subjectiveness*

People generally like to follow the opinion leaders' advises and take their suggestions to purchase a product and the visitors are more impressed by subjective comments. The all positive or all negative comments are hard to be trusted has been indicated by previous work [7]. However, this factor does not used to judge the positive or negative aspect but estimate the subjectiveness of bloggers' expression. As to the "subjective words", by definition of HowNet Knowledge Database, we take these 4,542 positive words and 4,333 negative words as our "subjective word set". The total occurrences of subjective word w_i in all the blog belonging to the blog post set Φ_b can be written as

$$O(b) = \sum_{w_j \in p_j} \sum_{p_j \in \Phi_b} occurrence(w_i, p_j),$$

where *occurrence* (w_i, p_j) represent the number of subjective term w_i occurring in blog post p_j. Let B is the set of all targeted blog sites. The subjectiveness score (s_b) of blog site b is formulated as:

$$s_b = O(b) / \max_{\beta \in B} \{O(\beta)\}.$$

- *Length of a blog post*

Align to the finding in Agarwal et al. [3], the length of a blog post is positively correlated with the number of comments which means longer posts attract people's attention. We also take the length of a blog post as a factor to determine marketing influence. The blog length score of (λ_b) of blog site b is computed by averaging the length of all the blogs in the blog post set Φ_b and expressed as:

$$\lambda_b = \left(\sum_{p_j \in \Phi_b} |p_j|\right) / |\Phi_b|,$$

where $|p_j|$ is the length of blog post p_j and $|\Phi_b|$ is the number of blogs written by blog site *b*.

- *Living time in network*

A blog site with longer living time may have more opportunities to interact with more other bloggers, thus the impact on the entire blog social network will be deeper and wider than the blog site which with shorter living time. Living time length score of (τ_b) of blog site b is calculated as:

$$\tau_b = T(b) / \max_{\beta \in B} \{T(\beta)\},$$

where $T(b)$ is the time interval between the post timestamps of first and the latest posted blogs in Φ_b.

- *CV (Content based Value) calculation*

To measure the content-based value of a blog site, we join subjectiveness (s), length of a blog post (λ) and living time in network (τ) together to enhance the MIV model. Then,

$$CV(b) = \tau_b(s_b + \lambda_b) .$$

3.3 MIV Value

Combining network-based and content-based value of a blog site, we could quantify the marketing influence value (MIV) of blog sites. Therefore marketers could use the MIV model to find out influential bloggers to run marketing campaigns. MIV model is developed by adding weighted network-based value (NV) and content-based value (CV).

$$MIV(b) = \omega_n NV(b) + \omega_c CV(b),$$

where $MIV(b)$ is the marketing influence value of blog site b , and ω_n and ω_c are the weights that can be used to adjust the contribution of network-based and content-based marketing influence value. A three-layer back-propagation neural network (BPNN) is employed to combine *NV*, *CV* and forecast the final MIV value. We collected 70 bloggers sampled from our crawled blogger set. Then, we invited 30 online users to score (0~100) the influential strength of the bloggers.

4. EXPERIMENTAL STUDY

In the following paragraph, we apply the proposed MIV model in an empirical study. We investigate the performance of different blog site selection algorithms and compare the results according different approaches.

4.1 Data Collection and Processes

We tested our model by using a dataset collected from Wretch blog [28, 30]. It is the largest weblog community in Taiwan with millions of users registered where users can upload photos to their albums, write the blog and interact with others. Since the purpose of our model is to find those nodes having power in influencing the buying decisions of others, Wretch/cate is especially suitable candidate to conduct the experiments. It provides abundant interactions (such as publish, citation, comment, friend and friend-of lists, etc.) about cate information. The dateset was collected at 14, May, 2008 under the category of cate. First, once the authors and related bloggers are crawled, we turn the focus on collecting articles and related information from their blog site. For the sake of computational efficiency, at most 10 articles for each target authors are crawled for further analysis. Table 1 lists the statistics about our experiment. Overall statistics illustrates the dimension of blogger (i.e. a blog site), and it reveals the characteristics of the overall data that we will use in our experiments.

Table 1. Statistics data

Statistics of our crawled blogger set	
Number of available bloggers	382
Number of available blog posts crawled	3,455
Average number of friend per blogger	13.312 /per blogger
Average number of friend-of per blogger	15.385 /per blogger
Average number of posts per blogger	180.796 /per blogger
Average live-time per blogger (day)	517.215 /per blogger
Average cumulative hotness per blog site	129948.259 /per blog site
Average words per post	409.683 /per post
Average number of comments per post	1.687 /per post
Average number of citations per post	0.005 /per post

4.2 Experimental Design

In this experiment, the default settings in our experiment assumed q value is 0.25 according to PageRank setting [29]. Fig 2 showed the calculation progress in our proposed MIV model. A neural network with a three-layer BPNN is developed to evaluate the overall influential strength. The *NV* and *CV* are the input neurons, and the MIV value is the output neuron. 50 neurons are set in hidden layer. The epochs for the ANN leaning us set be 500.

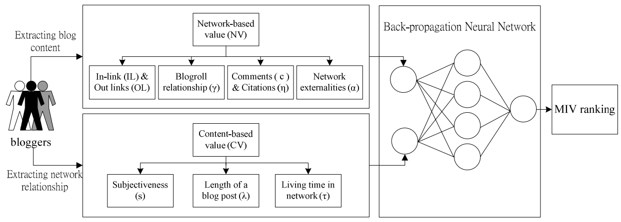

Figure 2. MIV calculation progress

4.3 Evaluation Results

We do experiments based on different selection algorithms and compare the results.; we compare the proposed approach with two other different blog site selection algorithms, such as out-degree centrality and betweenness centrality algorithm [9, 17]. The experimental results suggest a list of the top 20 bloggers with highest marketing influence value, which maximizes the marketing performance on this blog network as prior studies [6, 12, 13, 14] suggest that most users only access the documents/articles which are shown in the top-20 list.

- *Comparisons based on peer ratings*

It's important yet challenging to evaluate the impacts of chosen blog sites in the blog marketing domain. It's reasonable to assume that the explicit rating or ranking scores of bloggers on the website could be used to calculate the relevance between results of our proposed model and the real-world ranking mechanism. Thus, we built an evaluation process based on the Delphi method. 58 participants with background of e-commerce were invited for evaluating our MIV model. The evaluation process includes the following two stages. The detailed procedures are shown in Fig 3. In the first stage, the participants are invited to visit the blog sites on the recommendation list, mark the influential level and reported their comments. We asked them to evaluate the influence in 5 levels - "no influence," "below average," "average," "above average," and "high influence." In the second stage, the system lists the blog sites and these comments collected and the same participants re-visit the blog sites and make a second evaluation. Finally, the final group evaluation result is based on the majority of individual evaluation (voting).

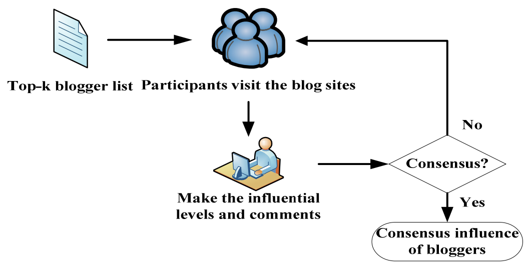

Figure 3. Consensus influence evaluation procedure

Table 2. Top 20 influential bloggers

Rank	Blogger#	MIV	Peer evaluation
1	Blogger361*	89.8188	High influence
2	blogger020*	88.1657	High influence
3	blogger337*	87.3116	Above average
4	blogger116*	87.1799	High influence
5	blogger287*	87.1787	Above average
6	blogger350*	86.9033	High influence
7	blogger216*	86.4317	Above average
8	blogger215*	85.6799	Above average
9	blogger009*	85.3481	High influence
10	blogger301*	84.5937	Above average
11	blogger195*	84.0265	High influence
12	blogger326*	82.8223	Average
13	blogger289	82.2988	No influence
14	blogger067*	81.4612	Above average
15	blogger085	80.1482	Below average
16	blogger194*	79.9163	Above average
17	blogger379*	78.7199	High influence
18	blogger320	78.7059	Below average
19	blogger104*	78.0841	Above average
20	blogger080*	77.4804	Average

The top 20 influential blogger list is shown in Table 2. A blogger with consensus peer evaluation "Average" or better grade is treated as an influential blogger. In Table 2, these bloggers are marked with asterisk (*). We can observe our recommendation gets approximately 85% accuracy according to user evaluations.

- *Comparisons based on official ratings*

Wretch has already identified 27 expert cate bloggers from their online users. There are 21 Wretch defined cate expert bloggers included in our blogger set. The bloggers who are in our recommendation list by italic format are the Wretch identified cate experts.

Table 3 compares the precision and recall values with respect to three approaches. Denote *N* as the number of official listed experts and *n* is the number of experts identified in top-K list for a discovering mechanism. The precision value is measured as *n*/*N* and recall value is calculated as *n*/*K*. We can observer that the performance of our proposed approach is significantly better than the other two. In addition, we also construct experiments to evaluate the effectiveness by changing the size of recommendation list. The 21 defined experts in Wretch are used as the recommended targets in following evaluation process. The precision and recall comparisons are shown in Figure 4 and Figure 5 respectively.

Table 3. Evaluation results based on official ratings

	Ours	Betweenness	Out-link
# of experts discovered	11	8	6
Precision value	55%	40%	30%
Recall value	44%	29.6%	23%

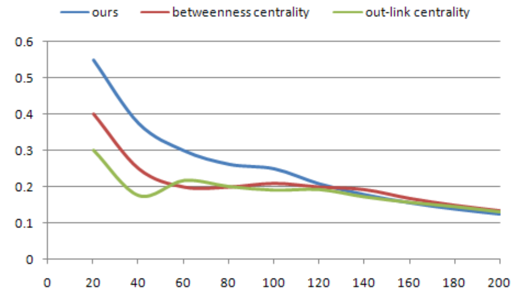

Figure 4. Precision comparisons

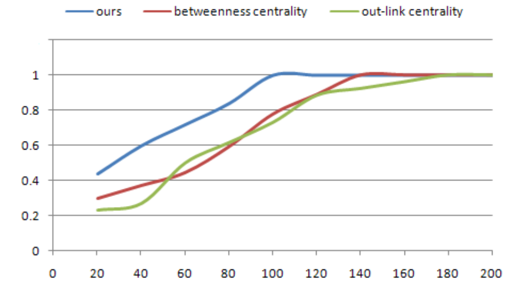

Figure 5. Recall comparisons

Our method could discover all experts in a top-100 list, while betweenness and out-link centrality discovered the experts in top-140 and 161 respectively. When the number of experts in the recommended list is more than top 120, these three lines will becomes very close to each other. It is because almost all the defined experts have already been discovered. In all situations, our proposed system significantly outperforms than others.

5. CONCLUSION AND FUTURE WORK

In this research, we construct a comprehensive model in which two dimensions of factors - network-based and content-based factors, are considered to measure the marketing influence strength and identify the potential authors in the blogosphere. The proposed approach can support the marketers/advertisers in promoting their products/services with less efforts and costs. Utilizing our proposed model, the most worthy bloggers with marketing value could be identified effectively. This study provides a feasible yet powerful way to generate a ranked list of bloggers according to their influential powers in improving the effectiveness of marketing activities. Our experimental resultsshows that this model could significantly reduce marketing cost and uncertainty and the proposed model have better performance than other approaches such betweenness and out-link metrics.

There are still a few of avenues for the future research. First, in this research, influence strength is measured based on general content and network characteristics. It is interesting to consider the domain in which the bloggers have powerful influences. Second, this research mainly focuses on the discovery of influential bloggers. The impact of blogging about the cate of specific food on the consumption behaviors can be further examined. Finally, in addition to the articles and network characteristics, the influential power of on-line reviews might consider other factors. For example, embedded photographs quality in the comments could be considered as one of the factors of content-based value.

6. ACKNOWLEDGEMENTS
This research was supported by the National Science Council of Taiwan (Republic of China) under the grant NSC 97-2410-H-009-035-MY2.

7. REFERENCES
[1] Abdul-Rahman, A. and Hailes, S. 2000. Supporting Trust in Virtual Communities. Proceedings of the 33rd Hawaii International Conference on System Sciences.

[2] Adar, E., Zhang, L., Adamic, L. and Lukose, R. 2004. Implicit structure and the dynamic of blogspace. Proceedings of the first Annual Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics.

[3] Agarwal, N., Liu, H., Tang, L. and Yu, P.S. 2008. Identifying the influential bloggers in a community. Proceedings of the international conference on Web search and web data mining.

[4] Brin, S. and Page, L. 1998. The anatomy of a large-scale hypertextual Web search engine. Computer Networks and ISDN Systems 30, 1-7, 107-117.

[5] Chiang, W.K., Zhang, D. and Zhou, L. 2006. Predicting and explaining patronage behavior toward web and traditional stores using neural networks: a comparative analysis with logistic regression. Decision Support Systems 41, 2, 514-531.

[6] Cacheda, F. and Viña, Á. 2001. Experiences retrieving information in the World Wide Web. Proceedings of Sixth IEEE Symposium on Computers and Communications.

[7] Ding, X., Liu, B. and Yu, P.S. 2008. A holistic lexicon-based approach to opinion mining. Proceedings of the international conference on Web search and web data mining.

[8] Domingos, P. and Richardson, M. 2001. Mining the network value of customers. Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining.

[9] Freemann, L.C., A set of measures of centrality based on betweenness. Sociometry 40, 35–41.

[10] Fujimura, K., Inoue, T. and Sugisaki, M. 2005. The EigenRumor Algorithm for Ranking Blogs. Proceedings of the second Annual Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics.

[11] Gill, K.E. 2004. How can we measure the influence of the blogosphere? WWW 2004 Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics.

[12] Hlscher, C. and Strube, G. 2000. Web search behavior of Internet experts and newbies. Proceedings of the 9th international World Wide Web conference on Computer networks : the international journal of computer and telecommunications netowrking.

[13] Jansen, B.J. and Spink, A. 2005. An analysis of web searching by European AlltheWeb.com users. Information Processing and Management 41, 2, 361-381.

[14] Jansen, M.B.J., Spink, A. and Saracevic, T. 2000. Real life, real users, and real needs: a study and analysis of user queries on the web. Information Processing and Management 36, 2, 207-227.

[15] Java, A., Kolari, P., Finin, T. and Oates, T. 2006. Modeling the spread of influence on the blogosphere. Proceedings of the 15th International World Wide Web Conference.

[16] Kempe, D., Kleinberg, J. and Tardos, E. 2003. Maximizing the spread of influence through a social network. Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining.

[17] Kiss, C. and Bichler, M. 2008. Identification of influencers - Measuring influence in customer networks. Decision Support Systems 46, 1, 233-253.

[18] Kleinberg, J.M. 1999. Authoritative sources in a hyperlinked environment. J. ACM 46, 5, 604-632.

[19] Kolari, P., Finin, T. and Lyons, K. 2007. On the structure, properties and utility of internal corporate blogs. International Conference on Weblogs and Social Media.

[20] Kritikopoulos, A., Sideri, M. and Varlamis, I. 2006. BlogRank: ranking weblogs based on connectivity and similarity features. Proceedings of the 2nd international workshop on Advanced architectures and algorithms for internet delivery and applications.

[21] Kuo R.J., Chen, J.A. 2004. A decision support system for order selection in electronic commerce based on fuzzy neural network supported by real-coded genetic algorithm. Expert Systems with Applications, 26, 2, 141-154.

[22] Li, Y.-M. and Chen, C.-W. 2009. A synthetical approach for blog recommendation: Combining trust, social relation, and semantic analysis. Expert System and Applications 36, 3, 6536-6547.

[23] Lin, Y.-R., Sundaram, H., Chi, Y., Tatemura, J. and Tseng, B.L. 2007. Blog Community Discovery and Evolution Based on Mutual Awareness Expansion. Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence.

[24] Richardson, M. and Domingos, P. 2002. Mining knowledge-sharing sites for viral marketing. Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining.

[25] Scoble, R. and Israel, S. 2006. Naked Conversations: How Blogs are changing the way business talk to their customers. John Wiley & Sons Inc.

[26] Subramani, M.R. and Rajagopalan, B. 2003. Knowledge-sharing and influence in online social networks via viral marketing. Communications of the ACM 46, 12, 300-307.

[27] Tajeddine, A., Kayssi, A., Chehab, A. and Artail, H. 2005. A Comprehensive Reputation-Based Trust Model for Distributed Systems. Workshop of the 1st International Conference on Security and Privacy for Emerging Areas in Communication Networks.

[28] Wikipedia Wretch. [http://en.wikipedia.org/wiki/Wretch_%28website%29]. Access on May 06, 2008.

[29] Wikipedia PageRank. [http://zh.wikipedia.org/wiki/PageRank]. Access on February 18, 2008.

[30] WretchBlog [http://www.wretch.cc/blog/]. Access on May 06, 2008.

[31] Zhan, J., Loh, H.T. and Liu, Y. 2009. Gather customer concerns from online product reviews - A text summarization approach. Expert System and Applications 36, 2, 2107-