Expert Systems with Applications 36 (2009) 6536–6547

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A synthetical approach for blog recommendation: Combining trust, social relation, and semantic analysis

Yung-Ming Li *, Ching-Wen Chen

Institute of Information Management, National Chiao Tung University, Hsinchu 300, Taiwan

ARTICLE INFO

Keywords:
Blogosphere
Trust model
Social networking
Information retrieval
Back-propagation neural network

ABSTRACT

Weblog is a good paradigm of online social network which constitutes web-based regularly updated journals with reverse chronological sequences of dated entries, usually with blogrolls on the sidebars, allowing bloggers link to favorite site which they are frequently visited. In this study we propose a blog recommendation mechanism that combines trust model, social relation and semantic analysis and illustrates how it can be applied to a prestigious online blogging system – wretch in Taiwan. By the results of experimental study, we found a number of implications from the Weblog network and several important theories in domain of social networking were empirically justified. The experimental evaluation reveals that the proposed recommendation mechanism is quite feasible and promising.

1. Introduction

Online social networking systems and peer-produced services have gained much attention as a social medium of viral marketing, which exploits existing social networks by inspiring bloggers to share their own posts or personal information with the other bloggers. The weblogs indeed provide 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 (Kolari, Finin, & Lyons, 2007). A blog social network has emerged as a powerful and potentially services-valued form of computer-mediated communication (CMC). More and more interactions take place in the blogosphere, combining the benefits of the accessibility of the web, the ease-of-use of interface and the incentive of blogging (i.e. share, recommend, comment…etc.). Blog becomes a viral marketing site based on peer-production and it is promoted yet induced by online person to person interactions. Moreover, there exists a large number of information in the blogosphere, including text-based blog entries (articles) and profile, pictures or figures and multimedia resources. This becomes problematic for users. How do they deal with information overload problems and how do they effectively retrieve information they consider important? This gives us an incentive to develop a blog recommender approach and design an information filtering mechanism.

A recommender system of weblog differs from others in several ways. First, recommendation target varies dramatically from product, movie, music, news, webpage, travel and tourism to all kinds of service, online auction seller and even virtual community (Lee, Ahn, & Han, 2007). It is important for us to find the characteristics of recommendation targets because the inappropriate use of recommendation may have a totally opposite effect by resulting unfavorable attitudes towards the recommendation target. Second, the blog recommender is also a provider. Unlike other contexts, blogs or bloggers in the entire blog network are highly dynamic and the recommendation environment changes fast. The blog recommendation mechanism must be more flexible and adaptable than the others. Third, it is more human-oriented. In other words, blog content itself is highly subjective and textual-sensitive for recommenders.

Blog search engine and blog recommender system serve similar function but differ to some extent. What is the difference between blog search engine and blog recommender system? This question emerges as the blog filtering approach such as search engine can also alleviate the mentioned problem. There are three folds of differences between them. First, information needs, real-time versus long-run. Some weblog aggregators e.g. Technorati provides tag-based search engine platform; moreover, Blogpulse and Daypop supply common keyword-based search engines just like Google and Yahoo but are applied in weblog domain. It allows users to find potential interesting postings, which many bloggers are talking and concerning about recently, with ease. In contrast to search engine technology, the proposed blog recommendation mechanism is long-run oriented. In other words, the former is popularization and the latter is more personalized. Second, pull versus push information: the former is a paradigm of technology of pull information. A search result is obtained as the query is submitted in advance. As for the latter, either pull or push technology could be employed

* Corresponding author. Tel.: +886 3 5712121; fax: +886 3 5723792.
E-mail addresses: yml@mail.nctu.edu.tw (Y.-M. Li), kisd0955@hotmail.com (C.-W. Chen).

0957-4174/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2008.07.077

to induce the recommendation result. Third, diversity of recommendation process: the former only considers the content and term comparison. As for the latter, it considers a multidimensional approaches and factors to implement the recommendation mechanism. In this study, the proposed mechanism takes three factors into consideration.

In blog recommendation context, it is important that how we introduce interesting, personalized and socially related weblogs of this peer-produced information to bloggers through recommendation mechanism. The objective of blog recommendation mechanism in this study is bloggers or blog posts (articles). Then what kinds of blog posts do we recommend? Is it most popular, or most trustworthy, or most similar in links or in blog content? These approaches and related researches inspire us to combine them to propose a synthetical recommendation mechanism in this study. We believe that trust model, social relation and semantic similarity play an important role in trust recommender system, social networking analysis and information retrieval/textual comparison, respectively and they are three crucial factors to help prepare the ground for the development of personalized and trustworthy recommendation mechanism.

The rest of paper is organized as follows: Section 2 presents related works. Section 3 designs a system framework of neural network-based recommendation mechanism. Section 4 elaborates on methodologies of trust model, social relation and semantic analysis. Section 5 proposes an experimental study to discover some characteristics of blog network and demonstrates the entire recommendation mechanism. Section 6 discusses the potential problems and some limitations. Section 7 concludes the paper.

2. Literature review

A fast-growing number of blog studies have shown that blog as social network can help researchers in understanding and analyzing certain implications and insights. It generated several issues and received lots of attention in several issues. The works (Adar, Zhang, Adamic, & Lukose, 2004; Fjimura et al., 2005; Kritikopoulos, Sideri, & Varlamis, 2006; Lin, Sundaram, Chi, Tatemura, & Tseng 2006) demonstrated social relation-based dimension to measure the importance and relationships of webpage or blog. The concept of blog ranking is similar to that of blog recommendation to some extent. (Fjimura et al., 2005) 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 (Brin & Page, 1998) and HITS (Keinberg, 1999) in that all are based on eigenvector calculation of the adjacency matrix of the links. However, the work in Kritikopoulos et al. (2006) 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. Works in Adar et al. (2004), and Kritikopoulos et al. (2006) 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. In our data set, only 57.22% of blog posts are isolated and without any comment and citation. The recommendation pool is large enough to perform our mechanism. Equally, in order to solve the sparsity problem, the extracted communities in Lin et al. (2006) only cover a portion of the entire blogosphere, the ranking method extract dense subgraphs from highly-ranked blogs.

Previous researchers suggested trust as another dimension to strengthen the reliability and robustness of recommender system. The works (Golbeck & Hendler, 2006; O'Donovan & Smyth, 2005; Papagelis, Plexousakis, & Kutsuras, 2005) applied trust to reinforce the ability of recommender system. Recommenders in blog network may have similar social relationships or contents to a target user (i.e. recommendation service requester) but they may not be a reliable predictor for inducing the recommendation. Using trust in recommender system will improve the ability of making accurate recommendation (O'Donovan & Smyth, 2005), which can solve partial weaknesses of traditional content-based, collaborative filtering (CF)-based recommendation approaches. In addition, (Golbeck & Hendler, 2006) use trust in recommender system to create predictive rating recommendations for movies. The accuracy of the trust-based predicted ratings for movies, compared with other approaches, is significantly better. Moreover, (Papagelis et al., 2005) proposes a trust-based method based on trust inferences to deal with the sparsity and the cold-start problems. Studies above have shown trust is critical when considering recommender system. Accordingly, our approach constructs a trust network by friend relationships where trust is central in these issues.

Semantic analysis is also important dimension to be taken into consideration. The works (Adar et al., 2004; Berendt & Navigli, 2006) indicated that applying semantic or textual-based analysis in blog domain is suitable and fruitful. Since the blog posts are strongly representative and we can discover the preferences and writing pattern of bloggers whom we want to recommend to. Traditional information retrieval (IR) technology is applied to handle the semantic of blog content. In examining the semantic similarity among weblogs, CKIP Chinese word segmentation system (Ma & Chen, 2003) helps us parse and stem the crawled post contents. Index terms are highlighted through IR/NLP approaches. Many syntax-based and semantics-based approaches exist to analyze the textual relationships among blogs (Tsai, Shih, & Chou, 2006). In Berendt and Navigli (2006), they proposed two methods for semantics-enhanced blogs analysis that allow the analyst to integrate domain-specific as well as general background knowledge. And the iRank in Adar et al. (2004) acts on implicit link structure to find blogs that initiate epidemics, which denote similarity between nodes in content and out-links. Undoubtedly, the content of blog post is an important source when inducing recommendation.

The researches (Song & Phoha, 2005; Weihua, Phoha, & Xu X., 2004) show that applying back propagation neural network in all kinds of domain of the learning ability to conduct forecast and prediction is appropriate. Under a multi-agent or peer-to-peer distributed environment, network is consisted of heterogeneous peers whose trust evaluation or rating standards may differ (Weihua et al., 2004). The issues are how to accurately and effectively predict trust value of an unknown party from multiple recommendations (Song & Phoha, 2005) by BPNN. However, in blog context, it may also help in deriving final recommendation score for each of the blog post which is expected to satisfy the user's preferences.

In this paper, we focus on the issues of combined trust model, social relation analysis, and semantic similarity as a means of recommending bloggers or blog posts. And the neural network plays is able to learn and capture the pattern of preferences of blog users and it is utilized to predict the final recommendation score of each blog post in our recommendation network.

3. System framework

3.1. Blog recommendation mechanisms

In this study, we propose an innovative weblog recommendation mechanism on the blogosphere which employs the trust model, social relation and semantic analysis to construct a more comprehensive and more personalized framework for each blogger on the entire blogspace. There are various important factors and dimensions we must take into consideration in blog recommendation context. We employ three underlying critical aspects of blogosphere, trustworthiness and reliability (TR), social intimacy and popularity (SIP) and semantic similarity (SS). Moreover, we present a neural network-based approach to learn and predict user's preference and affinity. By feeding these standardized scores into neural model, final recommendation score (FRS) of each blogger and blog post will be learned. Fig. 1 is the architecture of the proposed NN-based recommendation mechanism.

3.2. Trustworthiness and reliability (TR)

In literatures, very few literatures take trust among bloggers into consideration but it is widely used in social networking and distributed computing environment. By definition, trust degree in the social network connotes: belief and commitment. That is, as we said "A trusts B" which stands for A having a belief in B who will provide good opinions or behaviors in the future, and A is willing to accept it. In this study, we have the similar definition. A directed trust degree between bloggers A and B is a hybrid of referral trust and content-provision trust, a suitable representation of social relation between these bloggers in the blog network.

In blog recommendation context, every blogger has potential (resources) to be a provider. Such potential can be applied to every node in blog network. But in other domains, a specific query could not be satisfied by all agents but by resources providers. However, we denote it as concept of "resource specificity" for the reasons that every blogger in blog network could be recommender or recommendee for a given query, every available blogger could provide the contents (objects). In addition, users rely on the familiars' recommendation to retrieve certain information or services, since is more trustworthy and reliable because they have similar affinities and preferences.

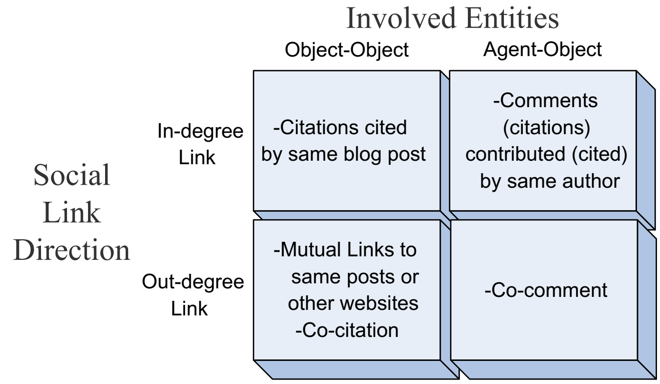

Fig. 2. Distribution of related social links among blogs.

3.3. Social intimacy and popularity (SIP)

SIP involves either explicit or implicit links which are used to represent the relationships and interconnections of the social network-based graphs. In blog linkage analysis context: nodes stand for bloggers or article posts and edges represent social or similarity relations between bloggers (article posts), which generally contain four social behaviors: comment, blogroll, citation, trackback (see ALi-Hasan and Adamic (2007) for more detail definitions) and virtually important similarity relationship is also a key factor. They reflect the social intimacy and interests relatedness between bloggers (posts) in the blog network.

Also the intrinsic sparsity problem of blog network is addressed, there are quite a few links among the entire blogspace and the majority of blogs are isolated. Hence, many approaches are introduced to tackle the problems by adding implicit links between blog entries, such as the similarity of counting the number of common tags/categories, the number of coupling URLs to news article and the number of authors posted in both weblogs (Kritikopoulos et al., 2006). We use the following two-dimension table to demonstrate crucial factors in inducing SIP score. Accordingly, the table

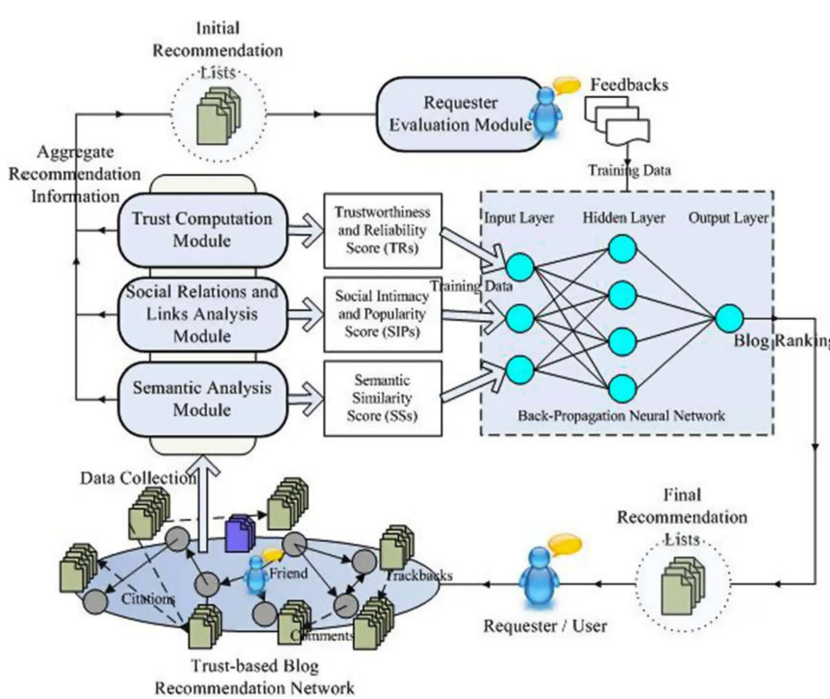

Fig. 1. The architecture of the proposed NN-based recommendation mechanism.

enables us to clarify the distribution of factors in computing the score of social intimacy similarity and helps us to reason and build our models base on these implicit links (see Fig. 2).

In this study, we utilize and consider the multiplicity of links which takes social intimacy and popularity as a basis to calculate our score of social relation in a more comprehensive and exquisite way.

3.4. Semantic similarity (SS)

The main purpose of semantic similarity analysis is to induce the most similar posts and to discover potentially interesting posts for the bloggers. However, blog search technology seems provide similar services, but actually it focuses on rating blog entries based on their similarity of posts contents or topics, once a set of keywords or tags is given. Compared with it, SS analysis provides full-text-based content matching approach to compute textual relatedness of a blog post pair. Therefore, information retrieval, text mining or social tagging methodology is proposed to handle these issues. As to this study, a traditional IR approach is applied to compute textual similarity between given weblog posts in entire blogspace.

Integrating the above concepts which rates the weblog posts according to its own trustworthiness and reliability, social intimacy and popularity, and semantic similarity score in a combinatorial manner, is able to induce a comprehensive and exquisite blog recommendation score. Then a weblog posts recommendation list is generated by ranking scores from high to low.

3.5. Neural network-based user evaluation process

In this study, a user is defined as a blogger who has his/her own weblog, interacting with other bloggers and has the needs to discover familiars in the blog social network.

In this step, a three-layer back-propagation neural network (BPNN) is employed to forecast the FRS (final recommendation score) for each weblog post after we get feedbacks from user evaluation results. The number of input neurons in the input layer is three (that is TR, SIP and SS score respectively). For the sake of output of training data of neural network, we design a web-based evaluation interface to collect the feedbacks from users according to the initial recommendation list. This process aims to generate the output data of back-propagation neural network for training, that is, the user feedback is deemed as the actual output of output layer in the back-propagation neural network.

To train the network, we set a threshold value as a performance target and train the network until the network reaches convergence. Then, the FRS can be derived for each bloggers/posts by employing the trained network. In other words, feed another TR, SIP and SS scores into trained network and FRS is generated in final. Due to the ability of learning and forecasting of BPNN, the proposed model could capture the social behaviors and preference patterns of users in which a truth-revealing recommendation result will be produced.

4. Research methodologies

This study proposes a neural network-based blog recommendation mechanism combined with the concepts of trust, social relation and semantic analysis. This mechanism contains the information of the blog network about trustworthiness and reliability, Social intimacy and popularity and Semantic similarity respectively. The whole process of recommendation mechanism is divided into several steps as shown in Fig. 3 and is described as the following sub-sections.

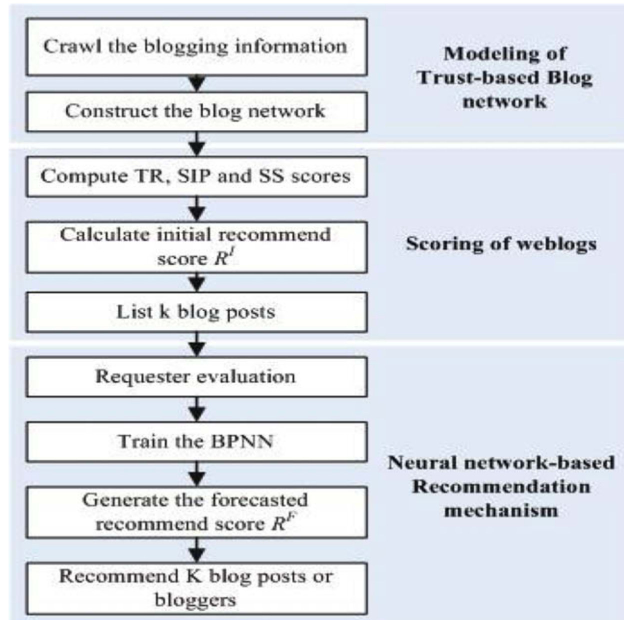

Fig. 3. The whole process of recommendation mechanism and its sub-sections.

Note that a recommendation score for an object (agent) in this study represents degrees of trustworthy, potentially alike in social interaction and semantically related in blog contents with respect to the recommendation service requester. In other words, the object (agent) with the higher score has more value and utility to recommend to requester, and he/she will have a greater preference and likeness toward the object (agent).

4.1. Trust-based blog network model

Crawl the blogging information. First of all, we take blogsite of requester as a starting point to search available and social-reachable agents i.e. recommenders, by performing search algorithm according to blogrolls on the side bar in the blogsite of each agent. These agents are connected level-by-level by friend or friend-of relationships in the blog network. Once the agents are decided and specified or the maximum number of searching level is reach, the members of the recommender are confirmed. Then we crawl blogging information (such as blog posts, hyperlinks, comments...etc) associate with each agent on the recommendation network.

Construct the blog network. To implement and evaluate the proposed model, we simulate a trust-based blog network which apply the concepts of agent and object in Fjimura et al. (2005). In this graph-based representation blog network (shown in Fig. 4), m agents (bloggers) and n objects (blog posts) are denoted as nodes and document-like icons, respectively. The relation edges in the network denote heterogeneous and multiplicity of links (whether explicit or implicit links), that is, it depends on the directions and entities involved here. Note that the constructed blog network forms and extends from the requester (node in yellow), then the trust information could be propagated and inferred in the agent layer. After that, the scope of object layer will be determined by these objects which can be reached by these agents in the agent layer. First of all, the problems of clarifying the existence of links and of classifying and annotating known links for both explicit and implicit ones are first steps toward identifying potential relationships in this incomplete graph. In this study, the relations are classified into following three aspects.

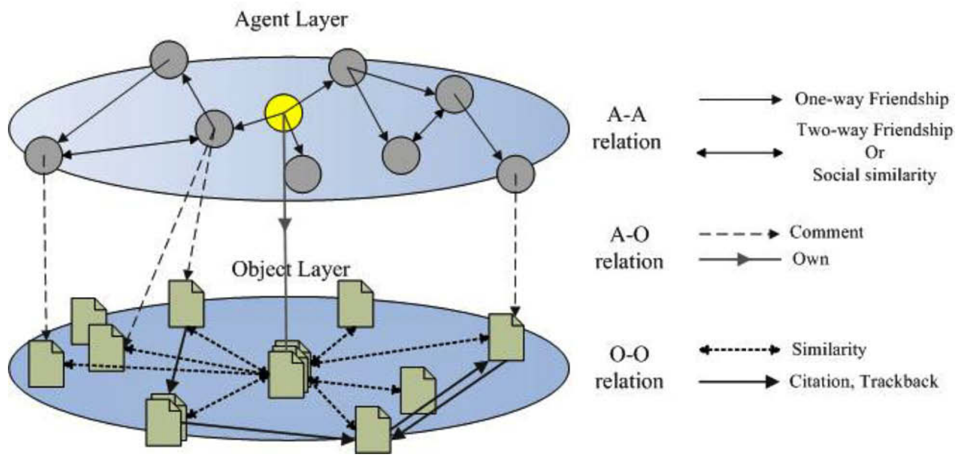

Fig. 4. The definitions and classifications of links among blog network.

4.1.1. Agent-to-agent relation (A–A relation)

A–A relation contains two kinds of relations. Firstly, a friend or friend-of relation, reflected in the blogroll, is a hyperlink from agent-to-agent. We quantify the relation as a degree of trustworthiness and reliability toward an agent who is worth to be conduct a belief and commitment that the agent will have a good referral or recommendation behaviors, i.e. trust degree. As a result, It forms (trust degree) the TR score.

Second sort of relation is about social similarity level which measures the strength of social intimacy and interaction in common between agents. In this section, not only real links in physical but also implicit similarity relations of social behaviors are taken into account, i.e. links in common, topic similarity, number of hyperlink in common, the number of same tags or comments contributed by same author in a post…etc. By aggregating these relations we could derive a social similarity score. In this study, we take behaviors of comment and citation for constructing a part of SIP score.

4.1.2. Agent-to-object relation (A–O relation)

In blog social networking environment, much of the interesting interaction occurs in comment behaviors and it is the most interactive and conversational way compared with other interactions. This kind of agent-to-object relation not only reveals the interests and social intimacy of blogger (commentator) toward specific blog post but also shows the popularity of bloggers. It is intuitive that a certain object will get a higher popularity score when it has more comments and citations (in-degree links) from other agents in spite of the community type, semantic of blog posts and recency/freshness factors. In examining the SIP score associated with popularity degree, comment is a crucial social behavior to express the social importance in blog network.

Another relation between agent and objects is possession relation and it implies that objects are submitted by an agent. Here is the entrance to connect agent with object layer for the purpose of inducing a personalized and requester-oriented social networking and computing mechanism.

4.1.3. Object to object relation (O–O relation)

Contributing to computation of SIP score, citation and trackback behaviors should be brought into model to improve the recommendation completeness, we especially emphasize on similarity between objects. Previous studies reckoned similarity as an important perspective in recommendation domain (Golbeck, 2006; Massa & Bhattacharjee, 2004; Matsuo & Yamamoto, 2007). In blog context, similarity plays the same role in recommending blog articles and bloggers. The proposed approach divides the concept of similarity into two sorts: Social intimacy similarity and semantic similarity of blog posts which associate with SIP and SS score, respectively.

4.2. Scoring approaches of weblogs

Calculate initial recommendation score R^{I} and list K blog posts. We compute a initial recommendation score (either for post or blogger) according to their scores of trustworthy, social relation and semantic similarity after a min–max standardization approach applied to each score (showed by upper case in Eq. (1)). An initial recommendation list is generated with a sequence of recommendation score ranking from high to low. Recommendation scores $R(i,j)$ for each post j of blogger i for given the requester r is defined as following:

$$R^{\mathrm{I}}(r,o_{ij}) = \alpha \mathrm{TR}^s(r,i) + \beta \mathrm{SIP}^s(r,o_{ij}) + \gamma \mathrm{SS}^s(r,o_{ij}), \quad (1)$$

where uppercase I of recommendation score R^{I} stands for initial recommendation score and uppercase s of TR, SIP and SS scores mean scores after the process of standardization. Parameters α, β and γ are the self-set weights of trust score, social relation score and semantic score of objects in the blog network respectively and the values are between 0 and 1.

Then the initial recommendation list, with top k R^{I} score and ranges from highest R^{I} score to lowest one, was induced for requester for further evaluation process. Each scoring approach is presented in the following three sub-sections.

4.2.1. Trust scores

The interpersonal trust values derive directly from blogroll relationships (i.e. the TR scores) in this study. All agents assign trust value to his/her friends listed in the blogroll on homepage of blog site. The computation of TR scores is divided into two steps: First, for a given requester (also blogger) r, we collect and aggregate trust information then form the trust-based blog network of him/her for further inference and filtering. Second, a search algorithm is applied to the constructed blog network in the former step, and set a maximum search layer as stopping criteria. The aim of this step is to find out social-reachable and available agents from the given requester who is the root of the blog network. These agents form the recommender set RC(r) of requester r. By listing in the following, the TR score of agent s is computed by trust inference mechanism and it is the most widely used one in trust-based social networking and computing approach (Golbeck, 2006):

$$\mathrm{TR}(r,s) = t_{rs} \text{ and}, t_{rs} = \frac{\sum_{j \in adj(r)} t_{rj} \times t_{js}}{\sum_{j \in adj(r)} t_{rj}} \tag{2}$$

where r is the requester of blog recommendation, s stands for these social-reachable and available agents, and $s \in \mathrm{RC}(r)$, t_{rs} is the value of trust degree from agent r to s, and $t_{is} \in [0,1]$, $adj(r)$ means adjacent agents of agent r, i.e. friends of blogger r.

4.2.2. Social relation scores

This section measures social intimacy and population (SIP) score of each agent in the blog network via their interrelationships and shared properties. Combining a complete view in recommendation process, SIP score is divided into SI and Popularity scores, SI addresses the social similarity strength or the degree of familiar on agent–agent aspect. While, Popularity emphasizes global reputation on object aspect (shown in Fig. 5).

SIP score is introduced in the following:

$$\mathrm{SIP}(r, O_{ij}) = \alpha \mathrm{SI}(r,i) + (1-\alpha)\mathrm{Popularity}(O_{ij}), \tag{3}$$

where $\mathrm{SIP}(r, O_{ij})$measures the scores of every object or agent in blog network given a requester agent r as a basis for comparison and computation, and $\mathrm{SI}(r,i)$ and $\mathrm{Popularity}(O_{ij})$ represents social intimacy relation and popularity scores respectively. α is the self-set weight.

Social intimacy captures the idea of social similarity by examining the degree of interaction between agents or of mutual behaviors (links) toward certain blogs or websites

$$\mathrm{SI}(r,i) = \mathrm{sim}(iL(r,A), iL(i,A)) + \mathrm{sim}(oL(r,A), oL(i,A)), \tag{4}$$

where r, i stands for the requester of blog recommendation (source agent) and certain agent respectively, and $r, i \in A$ A denotes a set of agents (or websites) which are social-reachable and available agents, i.e. agents (websites) which can be reached by links (hyperlinks) or inferences mechanism.$iL(r,A)$ is a vector which simply counts the number of social links from r to each of the agents in set A, where social links in here denote out-degree link which actually includes the situations of co-citation, co-comment and mutual link between the agents. $\mathrm{sim}(\cdot)$ is the function to compute the similarity between two agents by inner product calculation. Contrast to out-degree aspect, the latter part of formula measures the in-degree link which includes the situations of comments (citations) contributed (cited) by same author (blog post). However, $oL(r,A)$counts the number of social links from agent set A to agent r, which is shown in vector form.

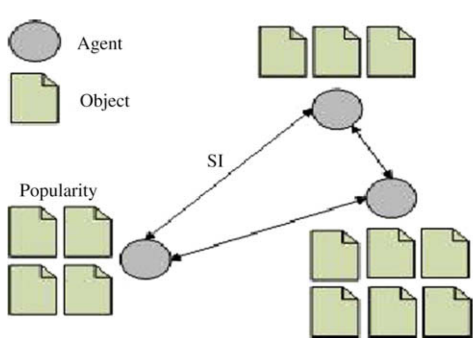

Fig. 5. The sketch map of social intimacy relation and popularity.

Popularity measures social importance of an agent or object in blog network. In general, three approaches are suitable for ranking nodes in a graph-based representation network, in-degree, HITS (Keinberg, 1999) and PageRank (Brin & Page, 1998). We measure the in-degree (the number of incoming links) in our model as a rough substitute for popularity for the ease of computing. Since an object u belonging to an agent s, we compute the aggregate value of u as a weighted sum of the relative number of comments and citations are as follows:

$$\mathrm{Popularity}(O_{ij}) = w_{co} \times \frac{\mathrm{Comment}(O_{ij})}{\max \mathrm{Comment}(A)} + w_{cl} \times \frac{\mathrm{Citation}(O_{ij})}{\max \mathrm{Citation}(A)}, \tag{5}$$

where $\mathrm{Comment}(O_{ij})(\mathrm{Citation}(O_{ij}))$ are the number of comments (citations) in object j of agent i. And $\max \mathrm{Comment}(A)(\max \mathrm{Citation}(A))$ is the maximum number of comments (citations) in our dataset. Obviously, the popularity score of an agent i, Popularity(i), is the sum of popularity score of objects belonging to i. The parameters w_{co} and w_{ci} are the weights of in-degree links from comment and citation behaviors respectively.

4.2.3. Semantic scores

Information retrieval techniques are originally used for extracting meaningful concepts and transforming unstructured text to structured data from documents. Especially in blog context, the recommendation target, source and the nature of interaction focus on texts, including topics, article contents which imply and convey significant information about the bloggers themselves. In this section, we apply traditional IR technique to compute the textual similarity among blogs and blog posts. There are several steps needed to calculate the semantic score of each post in the network (shown in Fig. 6).

Once the blogging data is crawled and HTML tags are removed, we apply CKIP (Chinese Knowledge and Information Processing) (Ma & Chem, 2003) Chinese word segmentation system to parse the content of blog post after the HTML tags are removed. While CKIP project in Academia Sinica proposes the Chinese parser to facilitate word segmentation and provides not only the functionality of word segmentation but also the morphological information of each word.

For the process of stop word removal, we extract several syntactical functions and morphological features (nouns and besides we select several kinds of verbs) that help us to extract useful terms for representing the documents. Then the remaining words are the index terms. A basic cosine similarity metric of term vectors with standard TFIDF (Manning & Schutze, 1999) weighting scheme is used to represent each index term of each blog article. Semantic score measures textual similarity of blog posts between requester and the other bloggers in the given blog network (once the blog network is constructed). Suppose there are n agents (bloggers) in the blog network. Semantic score is an agent-to-object score or object-to-object score and is defined as below:

$$\mathrm{SS}(r, O_{ij}) = \mathrm{sim}(q, d_{ij}), i \in j > 0 \text{ and } 0 \leqslant \mathrm{SS}(r, O_{ij}) \leqslant 1, \tag{6}$$

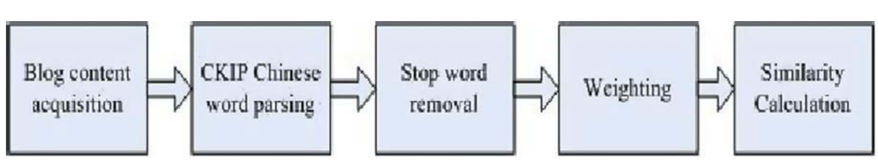

Fig. 6. The steps of semantics similarity analysis.

where q stands for index terms of blog postings which were published by requester r and we deem it as a query. Note that q could be generated by selecting any subset of objects of agent r. The variable d_{ij} is a vector of the TFIDF scores of index terms of blog post j of agent i.

The similarity comparison is limited within constructed blog network. On one hand, the agents in the blog network are introduced according to trust-based filtering mechanism which is more trustworthy and reliable to induce a better recommendation. On the other hand, the problem is computational efficiency which meant to deal with the problems of information overload and scale reduction of the blog recommendation source pool.

4.3. Neural network-based recommendation mechanism

A back-propagation neural network (BPNN) model is one of the most frequently used techniques for classification and prediction, and is special in accommodating complex and non-linear data relationships. Thus, in this section, BPNN is adopted to capture the implicit relationships between these factors (TR, SIP and SS) and requester's preferences in blog social network accurately in a comprehensive view to forecast the FRS for each object or agent.

Requester evaluation. Once the initial recommendation list of k blog posts (bloggers) is delivered to the requester, which accompanies with a detail principles of evaluation by a web-based interface to help user fill the form with the satisfaction scores with ease. For the requester, all he/she has to do is review these posts (bloggers) and make a unbiased evaluation by scoring each posts (bloggers) selected according to his/her own preference based on the degree of perceptibly relatedness and similarity with respect to himself/herself.

Train the BPNN. The characteristics, preference and social behaviors vary dramatically among human beings. Neural network-based recommendation mechanism is special for its leaning and forecasting ability to imply the implicit relationships behind these factors and requester's pattern of preference. Notably, a forecasted score for each object will be obtained and the weights of initial recommendation score with respect to three scores will be learned (i.e. weights α, β and γ for TR, SIP and SS scores respectively) through the neural network. Therefore, to train the back-propagation neural network, we combine three scores i.e. TR, SIP and SS, and the results from requester evaluation process as testing data for BPNN. Once the network is trained, it can be used to calculate the forecasted recommendation score R^{F} and then generate recommendation list of K blog posts or bloggers to the requester.

5. Experiment study

So far we have introduced trust model, social relation and semantic analysis into our model. They present crucial factors to guarantee high-quality recommendations in blog network. In this section, we apply the proposed recommendation framework to Wretch, a famous blog system accommodating millions of uses to interact with others (Wikipedia: http://en.wikipedia.org/wiki/Wretch_%28website%29) in Taiwan and show the entire recommendation processes. We then conduct an empirical experiment to examine the effectiveness of proposed blog recommendation mechanism, satisfaction level of service requester.

We begin by explaining how the dataset was collected. Then some statistical data will be presented such as the number of bloggers in the recommendation network, average number of friends of bloggers and of blog posts for each blogger. To follow, we introduce how to build trust network to calculate TR score. Experiment results and evaluations are addressed in the end.

5.1. Data descriptions

We describe our proposed mechanism by using a dataset collected from the Wretch (Wretch blog: http://www.wretch.cc/blog/) which is a Taiwanese community website. It is the most famous weblog community in Taiwan with millions of users registered now where users can upload photos to album, write the blog and interact with others by these services (Wikipedia: http://en.wikipedia.org/wiki/Wretch_%28website%29).

In early July 2007, we start crawl related blogging information including blogger account, friend relations, article id, article content (object), citations, comments and publish datetime for each blogger by using the crawler we designed, once the recommendation network is constructed. Note that, the objects are crawled according to the agents which have been crawled.

The detail statistics information of this experimental recommendation network is presented in Tables 1 and 2. It can be observed that the network size is drastically increasing, and we can predict that the network will achieve a saturated situation when the network spreads up to $5\sim6$ layer. That is, the network will be close to the entire blog network of Wretch (i.e. about 2.5 millions+ users).

An experimental small recommendation network about 20,000+ agents and 330,000+ objects will be constructed and limited the layer to 3rd layer, due to the reasons that the network size grows up exponentially with the layer increased, which will result in a decreasing computability of trust and semantic similarity.

To describe entire network, about 57.22% of objects are isolated and without any comment and citation. From (Fig. 7), we found that 99% of the objects have comments range from 0 to 15, 80% range from 0 to 2, but 57.4% of objects do not have any comments. Moreover, 99% of the objects do not have any citations. Because of the sparse nature of blogosphere we have mentioned earlier, our approach seeks to increase the density of the implicit links between bloggers and between blog posts. This will enhance the reliability and comprehensiveness of recommendation mechanism.

Notably, the recommendation network in this study is formed according to the requester's friend network i.e. trust network. In other words, we fetch the users, who are reachable walking the network of trust of starting requester, into our dataset. We conduct our experiments with pre-selected target requesters who provide recommendation information and evaluate the effectiveness of the proposed recommendation mechanism in this study.

Table 1
Statistics of recommendation network (up to 3rd layer)

Characteristics of recommendation network	Statistics
# of agent (blogger) in the network	22,336
# of object (blog post) in the network	338,614
Average # of friend of an agent	29.722
Average # of objects of an agent	15.160
Average # of comments of an object	2.382
Average # of citations of an object	0.084

Table 2
The # of agent and friend relationship in each layer according to the root: "chiang1000"

#/layer	Root	1st layer	2nd layer	3rd layer	4th layer
The # of agent	1	23	927	21,384	299,539
The # of friend relationship	23	972	30,299	632,389	NA

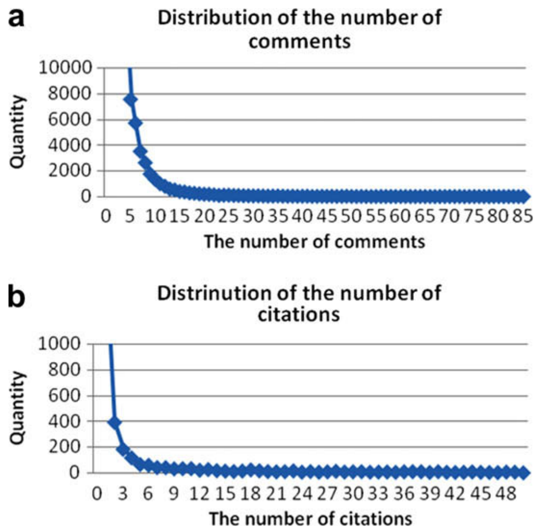

Fig. 7. The distributions of the number of (a) comments and (b) citations in our dataset.

5.2. Building trust network

Past works took user profile similarity or rating similarity over items as degree of trust among users (Golbeck, 2006; Golbeck & Hendler, 2006). However, in the context of blog recommendation, these methods cannot not be used due to the sparsity problem. As such, we need to develop a trust-generating network suitable in the blog environment. See Fig. 8 for a visualization of trust network, taken from one of Wretch blogger account "chiang1000". A complete line represents a trust value and a dotted line means lack of trust value toward certain blogger. The direction of arrow stands for the direction of trust information.

Trust can be generated by having trust values directly assigned by each blogger to his/her friends. We design an interface to receive trust from blogger. When using the interface, bloggers are asked to assign a trust value to each of his/her friends on the blogroll of blogsite. Being bothering for bloggers, this is one of ways to realistically capture the degree of trust. Then, we calculate trust value of each blogger for requester either by trust inference approach or averaging trust value toward certain blogger once the inference approach can't reach it. As to the rest of bloggers who lack of trust information in this recommendation network, we then ignore them.

5.3. Experiment results and evaluations

The experiment is conducted with 6 Wretch bloggers who did not have prior knowledge about the recommendation algorithms used in the system and who had different preferences. The target users are asked to examine the initial recommendation list to judge the recommendation results on a 10-point scale ranging from strongly satisfaction to strongly dissatisfaction (1, very unsatisfied; 5, average; 10, very satisfied). The averaged satisfaction score can be used to indicate the degree of fitness and user satisfaction between the users' preferences and recommended articles or bloggers.

5.3.1. Recommendation strategies

We design seven different recommendation strategies to evaluate the proposed mechanism, some of which are commonly used approaches provided by blog service providers (BSP) (i.e., wretch) (Wretch blog: http://www.wretch.cc/blog/) as the comparison benchmarks. The followings are the different recommendation strategies:

1. ANN + ALL, which is the approach proposed in this study. We Apply back-propagation neuron network to learn the final recommendation scores from the combinations of TR, SIP and SS scores.

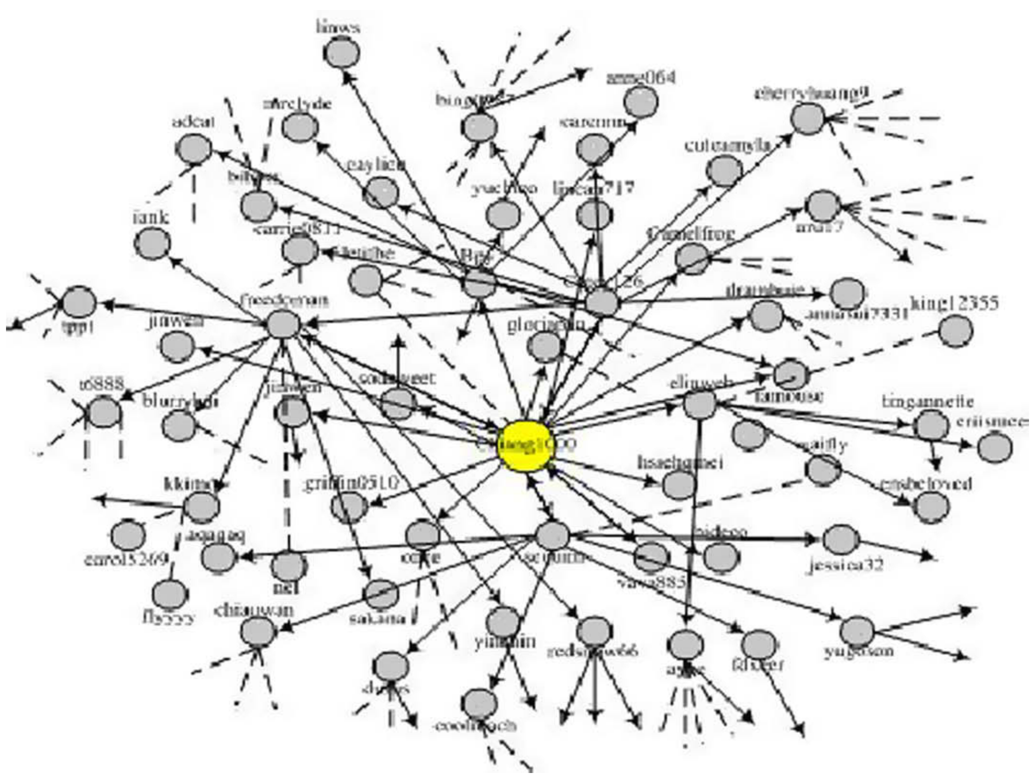

Fig. 8. A visualization of portion of trust network which is spanned from the core of blogger account "chiang1000".

2. ALL, which results in initial recommendation scores. Without BPNN to learn the non-linear relationships between TR, SIP, SS scores and final recommendation scores is addressed. In other words, recommendation scores is formed by weighted sum of these scores (see Eq. (1)). In this study, we set $\alpha = 0.3$, $\beta = 0.3$ and $\gamma = 0.4$.
3. SS, which is purely taken semantic similarity of contents into consideration. In this strategy, we ask six target users to select an article published in their blog site. We then focus on processing this selected article to calculate the content similarity with other articles in recommendation pool. It is similar to "full-text search" in a different form to some extent.
4. Random, which simply recommends articles or bloggers at random.
5. Comment, which recommends top-n articles or bloggers with more numbers of comments at certain time period.
6. Citation, which recommends top-n articles or bloggers with more numbers of citations at certain time period.
7. Hotness, which recommends top-n hottest articles or bloggers. The degree of hotness is dependent on the number of visitors of blogsite at certain time period.

In this experiment, we want to emphasize the power and robustness of hybridization of these factors accompanied with the preference predicting ability of BPNN in recommending weblogs (i.e., ANN + ALL strategy).

5.3.2. Neural network prediction model

We apply back propagation neural network (BPNN) to recognize the preference patterns and predict the final recommendation score of each target user in our mechanism. We utilized neural network toolbox of Matlab software to implement our model. The following table is the relevant network parameters and learning settings used in this experiment (Table 3).

We apply adaptive learning rate approach to accelerate the convergence of back-propagation learning to adjust the learning rate parameter during learning. Learning rate (lr) is multiplied by parameter lr_inc (lr_dec) whenever the performance function has an incremental increase (reduced).

A total of 20 subjects of each target users are gathered from the requester evaluation, and they are divided into 80%/20% training/testing data. Thus, there are 16 training subjects used for BPNN and 4 testing subjects for evaluation of the prediction ability of BPNN model. Therefore, we applied BPNN to select the better training parameters for the generating final recommendation list. The mean absolute prediction error (MAPE) and the root mean square error (RMSE) are adopted to evaluate the BPNN effectiveness. The according formula is shown in Eqs. (7) and (8)

$$\mathrm{MAPE} = \frac{1}{n}\Sigma_{i=1}^{n}\left|\frac{y_i - x_i}{x_i}\right| \times 100\%, \tag{7}$$

$$\mathrm{RMSE} = \sqrt{\frac{\Sigma_{i=1}^{n}(y_i - x_i)^2}{n}} \quad i = 1, \ldots, n, \tag{8}$$

Table 3
Network parameters and learning settings

Parameters	Value
Initial learning rate (lr)	0.001
lr_inc	0.1
lr_dec	10
Epochs	500
Number of hidden layer	1
Number of input neurons	3
Number of hidden neurons	5/10/15/20

where y_i is the predicted output, x_i is the actual output and n is the number of tested data.

When the MAPE and RMSE of test data set is more close to 0, indicate that BPNN model has more precise prediction ability.

In training parameter settings, 5, 10, 15 and 20 units of neurons in hidden layer are evaluated individually to derive a better BPNN model to minimize the error function for each user. Since the preference patterns are different, prediction models vary with users. In line with the concept, we develop different BPNN models for different users. The evaluation results of recommending articles and bloggers, displayed in average RMSE (MAPE%), of each BPNN models under different number of neurons in hidden layer of each user are listed in Tables 4 and 5. For each score listed in table is generated from taking average of five different trials with same parameters and setting.

The smallest RMSE value was marked in bold face in each row to denote better prediction ability of the BPNN model in both tables. This allows us to choose the appropriate hidden neuron number. We can observe that the MAPE does not perform well (significantly low). This may be due to the reason that training data is not insufficient enough to capture the complicated human decision patterns. The MAPE varied with different users, ranging from 10% more to 70% more in average. Which shown that the preference patterns of each user is rather different and hard to capture if training data is rather small. However, one may gives totally opposite satisfied scores at different time. As such, the prediction model should keep learning and adaptive to user's variability by feeding more and more training data. Although the overall average RMSE (MAPE) taken for predicting final recommendation scores of articles and bloggers is 0.199 (31.03%) and 0.329 (22.435%) respectively, the proposed mechanism still outperform than the others.

Table 4
The evaluation results of recommending articles, average RMSE (MAPE%) of each BPNN models under different number of neurons in hidden layer of each user

User (account_id) / # of neurons in hidden layer	5	10	15	20
1 (Chiang1000)	0.31 (35.571)	0.366 (37.242)	0.281 (36.131)	0.38 (40.875)
2 (Freedoman)	0.087 (55.394)	0.079 (47.962)	0.083 (54.88)	0.085 (54.979)
3 (Cutey126)	0.176 (24.01)	0.27 (37.281)	0.225 (30.901)	0.22 (30.187)
4 (Vivachu)	0.234 (23.808)	0.246 (25.973)	0.237 (24.83)	0.252 (26.806)
5 (Vava885)	0.277 (27.206)	0.259 (27.988)	0.369 (38.826)	0.286 (29.49)
6 (Anny0307)	0.165 (27.621)	0.519 (89.643)	0.437 (75.116)	0.406 (66.79)

Table 5
The evaluation results of recommending bloggers, average RMSE (MAPE%) of each BPNN models under different number of neurons in hidden layer of each user

User (account_id) / # of neurons in hidden layer	5	10	15	20
1 (Chiang1000)	0.335 (140.429)	0.158 (39.699)	0.459 (87.63)	0.13 (46.17)
2 (Freedoman)	0.142 (12.624)	0.112 (10.693)	0.164 (13.765)	0.104 (8.338)
3 (Cutey126)	0.295 (31.287)	0.253 (25.937)	0.174 (19.229)	0.18 (17.476)
4 (Vivachu)	0.25 (29.507)	0.17 (19.845)	0.481 (57.765)	0.364 (42.698)
5 (Vava885)	0.135 (18.552)	0.249 (32.192)	0.224 (29.106)	0.188 (25.121)

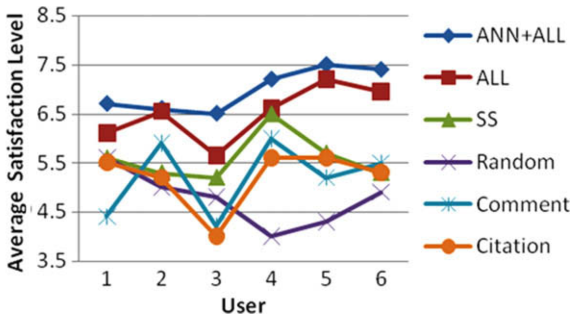

Fig. 9. The evaluation results of recommending articles.

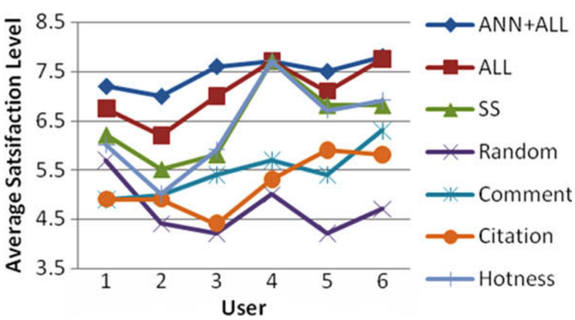

Fig. 10. The evaluation results of recommending bloggers.

5.3.3. User evaluation results

The following figures indicate the strategy of ANN + ALL and ALL being the best and second best respectively among other strategies. (Figs. 9 and 10) confirm the proposed blog recommendation mechanism is the best in average satisfaction level, compared to other approaches.

The statistical test (e.g. paired sample *t*-test) is used to further confirm the significance of the differences in the recommendation results. As shown in Table 6–16, at 95% significant level, both results of recommending articles and bloggers are statistically significant in terms of average satisfaction level. The results reveal that the proposed synthetical neural network-based approach is the best compared to others in the domain of blog recommendation.

6. Discussions

Our mechanism, combined TR, SIP and SS, showed to be an effective in recommending blog articles or bloggers to users. An

Table 6
The statistical verification results of recommending articles: ANN + ALL versus ALL

	Number of users	Mean	Std. dev.	*T*-value	$t_{0.05,5}$
ALL + ALL	6	6.983	0.436		
ALL	6	6.508	0.563	4.2	2.015
Paired difference	6	0.475	0.277		

Table 7
The statistical verification results of recommending articles: ANN + ALL versus SS

	Number of users	Mean	Std. dev.	*T*-value	$t_{0.05,5}$
ALL + ALL	6	6.983	0.436		
ALL	6	5.6	0.482	6.781	2.015
Paired difference	6	1.383	0.5		

Table 8
The statistical verification results of recommending articles: ANN + ALL versus Random

	Number of users	Mean	Std. dev.	*T*-value	$t_{0.05,5}$
ALL + ALL	6	6.983	0.436		
ALL	6	4.767	0.561	6.141	2.015
Paired difference	6	2.216	0.884		

Table 9
The statistical verification results of recommending articles: ANN + ALL versus Comment

	Number of users	Mean	Std. dev.	*T*-value	$t_{0.05,5}$
ALL + ALL	6	6.983	0.436		
ALL	6	5.2	0.756	6.401	2.015
Paired difference	6	1.783	0.682		

Table 10
The statistical verification results of recommending articles: ANN + ALL versus Citation

	Number of users	Mean	Std. dev.	*T*-value	$t_{0.05,5}$
ALL + ALL	6	6.983	0.436		
ALL	6	5.2	0.61	9.115	2.015
Paired difference	6	1.783	0.479		

Table 11
The statistical verification results of recommending bloggers: ANN + ALL versus ALL

	Number of users	Mean	Std. dev.	*T*-value	$t_{0.05,5}$
ALL + ALL	6	7.467	0.308		
ALL	6	7.083	0.587	3.02	2.015
Paired difference	6	0.384	0.311		

Table 12
The statistical verification results of recommending bloggers: ANN + ALL versus SS

	Number of users	Mean	Std. dev.	*T*-value	$t_{0.05,5}$
ALL + ALL	6	7.467	0.308		
ALL	6	6.467	0.799	3.893	2.015
Paired difference	6	1.0	0.629		

Table 13
The statistical verification results of recommending bloggers: ANN + ALL versus Random

	Number of users	Mean	Std. dev.	*T*-value	$t_{0.05,5}$
ALL + ALL	6	7.467	0.308		
ALL	6	4.7	0.58	9.714	2.015
Paired difference	6	2.767	0.698		

Table 14
The statistical verification results of recommending bloggers: ANN + ALL versus Comment

	Number of users	Mean	Std. dev.	*T*-value	$t_{0.05,5}$
ALL + ALL	6	7.467	0.308		
ALL	6	5.45	0.509	17.725	2.015
Paired difference	6	2.017	0.279		

Table 15
The statistical verification results of recommending bloggers: ANN + ALL versus Citation

	Number of users	Mean	Std. dev.	*T*-value	$t_{0.05,5}$
ALL + ALL	6	7.467	0.308		
ALL	6	5.2	0.58	10.37	2.015
Paired difference	6	2.267	0.535		

Table 16
The statistical verification results of recommending bloggers: ANN + ALL versus Hotness

	Number of users	Mean	Std. dev.	*T*-value	$t_{0.05,5}$
ALL + ALL	6	7.467	0.308		
ALL	6	6.367	0.937	3.795	2.015
Paired difference	6	1.1	0.71		

experimental study is shown how these components combined together will induce the final recommendation score. The trust models defined in this work can not only be used to enhance recommendation trustworthiness and reliability but also be utilized to increase the robustness of CF-based recommendation systems (O'Donovan, & Smyth, 2005). However, the information related to trust degree is not available if we utilize real data from online blogging system, it means existing online blogging system does not contain the concept of trust degree between any pair of friend relationship.

There are some limitations in this work. First, to capture trust information in the real world, we quantify it by asking users to assign the trust values. The invasive requirements toward users thus may cause some disfavor and the trustworthy issues (i.e. some misleading or skewed situations of recommender system). Obviously, the phenomenon where over than half of objects are isolated, will debase the value of SIP score. This may causes the recommendation score lays particular stress on the other two scores (i.e. TR and SS), which distort the recommendation scores and denotation of recommendation mechanism. Second, our trust models are constructed on an agent-to-agent level which cannot reflect trustworthiness in an object to object level. That is, with regard to objects of certain agent, we treat each object as the same trust level and each of them has the same trust value relative to the requester. In the future work, we will design a more comprehensive trust model to tackle with this issue to induce a complete and robust recommendation mechanism.

As to SS score, we design an interface for the requester to select some of his/her posts to compare content similarity with others. Unlike search engine, some brief keywords would induce numbers of results which make users hard to digest and unable to find what they really want. More index terms would be helpful for users to accurately locate the needed information (Yang, Yu, Valerio, Zhang, & Ke, 2007). In this work, article selecting process (i.e. select the posts to list into comparison target) would indeed increase the efficiency and accuracy of calculation of semantic similarity. As for processing procedures of Chinese words, each step could be refined and advanced for more accuracy calculation of SS scores.

7. Conclusions

This paper proposes a synthetical blog recommendation mechanism to personally recommend suitable blog articles or bloggers to users in blogosphere of practice. We have combined trust model, social relation and semantic analysis to develop our model and illustrated how it can be applied to a prestigious online blogging system – Wretch in Taiwan. Trust model measures the trustworthiness and reliability of the targets, social relation addresses the social intimacy and similarity of social behaviors in blog social network where both explicit and implicit links are considered. Semantic analysis simply compares the textual similarity of blog articles.

Major findings from the evaluation of the proposed blog recommendation mechanism are summarized as follows. In constructing recommendation network, we found that the recommendation network will almost contains majority of bloggers of Wretch when the network spreads up to 5–6 layer. The network dramatically grows with about 25.1 times in average with the increase of spreading layer. Although the network becomes more and more saturated, the expanding scope will more and more convergent until all of the interconnected bloggers are in the network. This "small-world" of blog recommendation network thus reveals the well known theory of "six-degrees". That is, most bloggers in recommendation network can be linked on average six degrees of traversal, except for isolated bloggers. From the experimental evaluation results, we can observe that the MAPE performance of the model seems insignificant under the limitation of the insufficient training data, even though the recommendation prediction model still outperformed than the other approaches, including the synthetical approach without BPNN training.

Acknowledgement

This research was supported by the National Science Council of Taiwan (Republic of China) under the Grant NSC 96-2416-H-009-017.

References

Adar, E., Zhang, L., Adamic, L., & Lukose, R. (2004). Implicit structure and the dynamics of blogspace. In *workshop on the weblogging ecosystem: aggregation, analysis and dynamics, WWW*.
ALi-Hasan, N., & Adamic, L. (2007). Expressing social relationships on the blog through links and comments. *ICWSM*.
Berendt, B., & Navigli, R. (2006). Finding your way through blogspace: Using semantics for cross-domain blog analysis. *In AAAI symposium on computational approaches to analyzing weblogs*.
Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *In proceedings of seventh international World Wide Web conference*.
Fujimura, K., Inoue, T., & Sugisaki, M. (2005). The EigenRumer algorithm for ranking blogs. In *second annual workshop on the weblogging ecosystem: aggregation, analysis and dynamics, WWW*.
Golbeck, J. (2006). Trust and nuanced profile similarity in online social networks. *Journal of Artificial Intelligence Research*.
Golbeck, J., & Hendler, J. (2006). FilmTrust: Movie recommendations using trust in web-based social networks. *IEEE Consumer Communications and Networking Conference*.
Keinberg, J. M. (1999). Authoritative sources in hyperlinked environment.. *Proceedings of the ninth annual ACM-SIAM symposium on discrete algorithms, 46*(5).
Kolari, P., Finin, T., & Lyons, K. (2007). On the structure, properties and utility of internal corporate blogs. *ICWSM*.
Kritikopoulos, A., Sideri, M., & Varlamis, I. (2006). BlogRank: Ranking weblogs based on connectivity and similarity features. In *Proceedings of the second international workshop on advanced architectures and algorithms for internet delivery and applications* (p. 198).
Lee, H.-Y., Ahn, H., & Han, I. (2007). VCR: Virtual community recommender using the technology acceptance model and the user's needs type. *Expert Systems with Applications, 33*(4), 984–995. 12.
Lin, Y. -R., Sundaram, H., Chi, Y., Tatemura, J., & Tseng, B. (2006). Discovery of blog communities based on mutual awareness. In *Third annual workshop on the weblogging ecosystem*.
Ma, W. -Y., & Chen, K. -J. (2003). Introduction to CKIP Chinese word segmentation system for the first international Chinese word segmentation bakeoff. In *Proceedings of the Second SIGHAN Workshop on Chinese Language Processing* (pp. 168–171). <http://ckipsvr.iis.sinica.edu.tw/>.
Manning, C. D., & Schutze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.
Massa, P., & Bhattacharjee, B. (2004). Using trust in recommender systems: an experimental analysis. In *Proceedings of second international conference on trust management*.

Matsuo, Y., & Yamamoto, H. (2007). Diffusion of recommendation through a trust network. *ICWSM*.

O'Donovan, J., & Smyth, B. (2005). Trust in recommender systems. In *Proceedings of the 10th international conference on intelligent user interfaces* (pp. 167–174).

Papagelis, M., Plexousakis, D., & Kutsuras, T. (2005). Alleviating the sparsity problem of collaborative filtering using trust inference. In *Proceedings of iTrust* (pp. 224–239).

Song, W., & Phoha, V. V. (2005). Opinion filtered recommendation trust model in peer-to-peer networks. *Lecture Notes in Computer Science*, 237–244.

Tsai, T.-M., Shih, C.-C., & Chou, S.-C. T. (2006). Personalized blog recommendation: Using the value, semantic, and social model. *Innovations in Information Technology*, 1–5.

Weihua, S., Phoha, V. V, & Xu X. (2004). An adaptive recommendation trust model in multiagent system. In *Proceedings of the IEEE/WIC/ACM international conference on intelligent agent technology (IAT' 04)* (Vol. 9, pp. 462–465).

Yang, K., Yu, N., Valerio, A., Zhang, H., & Ke, W. (2007). Fusion approach to finding opinions in blogosphere. *ICWSM*.