Reputation Based Academic Evaluation in a Research Platform

Kun Yu
Department of Computer Engineering, Huaiyin Institute of Technology, Huaian, Jiangsu, China
Email: varguard@yeah.net

Jianhong Chen
Department of Computer Engineering, Huaiyin Institute of Technology, Huaian, Jiangsu, China
Email: chenjianhong@hyit.edu.cn

Abstract—**Researchers have to face with huge information in their daily works. It is hard for them to screening for valuable information from huge volume of data. Reputation of literatures, publications, or scholars can help the researches to relieve their puzzle and advance their research ability. In this paper, the problem of screening is presented in a realized research platform. Reputation is modeled by synthesizing four elements: literature, author, source and reader. The perceptible interactions, such as reference, comment and P2P communication is considered to be the relationships between each pair of elements and help improve the accuracy of reputation. The reputation we build is similar to impact factor and PageRank, but it is more complex and is expected to be more robust in realistic environments, which has been proved by simulations. An iterative algorithm is introduced to evaluate reputation in a distributed mode. Simulations prove the practicality and effectiveness of the scheme we have proposed.**

Index Terms—**academic evaluating; reputation model; impact factor**

I. Introduction

Recently researchers can benefit from remarkable development of computer science and communication technology, and must face with new problem brought about by information explosion. It involves in the rapid increase of papers, web pages and other new media. Science research becomes more convenient but and more difficult. How to find the most appropriate information is becoming a key factor for the success of research[1].

Traditionally, face-to-face interactions are the most common and believable way to gain knowledge and clues for next step of research. Direct talk plays the role of content sifter. But the speed of knowledge propagation is slow and limits the efficiency of research output. New computer and communication technologies, such as BBS, Email, instant communication and WWW, facilitate knowledge acquisition in modern studies. A main way to gain new knowledge in academic research is through reading scientific documents, which can be easily got from online databases. In the meantime, researchers can exchange their ideas through Email, BBS or other internet tools. It may demonstrate great advantage over traditional methods if researchers can easily get the information they indeed need.

Reputation or rank is usually efficient to help them to choice the most suitable content. Among the applications of this kind, impact factor[2] and PageRank[3] are maybe the most famous two, which have been used widely. Not by chance, both them use a similar algorithm that regards the relationship between two elements as a recommendation which finally contributes to reputation of the receptor. In impact factor algorithm, references play this role and in PageRank, hyperlinks do so. However, in both reputation models, recommendations is only belong to single class, that is to say, each recommendation has only its weight, but without any difference in importance. For example in impact factor, the citation of a paper brings a weighted recommendation to the paper that only depend on the IF of the journal the paper is published on. As we will see in next section, it is not always true when more factors are taken into considered. In those cases, the recommendation is not only about the importance of the referrer, but the class of the recommendation. The multi-dimensional reputation makes it possible to introduce more clues to evaluate reputation more exactly.

In this paper, we developed a cooperative research platform, CRP. The platform helps researchers and learners to meet their need with the least effort. The system includes a research forum and a P2P communication tool. Literature indexes and user comments are the main part of the research forum. And P2P tool can facilitate the exchange of ideas and reviews about a paper or a research work. So there are many clues of the relationship of any two elements, such as comment, access times and friend list in P2P tool. Friend list is extra important because it implies the evaluation directly to someone else without the shortage of the bias of review articles, and includes some factors that does not directly

doi:10.4304/jsw.7.12.2749-2754

affect academic value, but are about interest similarity, reliability, participation degree, and incentive mechanism.

Searching is a key function of the platform, including paper search and user search. It combines keyword search with reputation to sort information by quality. A user can also searches for men who are suitable for his research field or his unsolved problem. Reciprocity and trustworthiness are also considered. The reputation is used to filter the users who cannot be trusted enough. The users with similar interests can be located based on their behaviors and the relationship network. In fact, searching provides an incentive method: the user with higher reputation will have more chance to take part in interaction and have more chance to be helped, so he is willing to contribute more to the system.

II. BACKGROUND

The Reputation is a metric of entity quality, which can only be evaluated by others who once interacted with this entity, directly or indirectly. The indirect interact is called recommendation. In academic research, the most representative reputation is SCI[2]. However, in calculate SCI, only periodicals and references are considered which can't distinguish different literatures of a same journal[4]. The distribution of IF isn’t even, so paper in some fields cannot have high IF that cumber the comparison between different domains[5]. In the past few years, several new measures has been put forward which provide the readers more choices than SCI to evaluate a paper [6].

Usually, academic reputation are not judged based on single quantitative variable but through synthesizing multiple subjective indexes. The Index Copernicus Scientists[7] besides providing scientists with global scientists networking and international research collaboration, present a multi-parameter career assessment system which analyses the researcher individual profile. Journal to Field Impact Score (JFIS) [8] developed an alternative system for the journal impact evaluation. Its source to compute index includes literatures, technical reports, notes and reviews. With extended data source, Castelnuovo focused on the reputation of simple researcher, which is called Single Researcher Impact Factor[9].

PageRank [3,10] is a more efficient reputation evaluation algorithm which is initially used to web page ranking. Compared to impact factor which only takes citation times into considered, PageRank extends impact factor by introducing weighted links to improve the validity of the information recommendation[11]. Google Scholar is special application of PageRank in the academic field with a broader range of open data sources: books, technical reports etc. Popular journals such as review journals with little prestige could have a very high IF and a very low weighted PageRank. EngerTrust[12] models reputation by the concept of recommendation which is similar to hyperlink and citation. EngerTrust also weighs recommendation based on recommending credibility like the weighted citation analysis. Recently some new methods towards the reputation or trust model have proposed too[13][14].

Recommender system[15] is a useful tool to help user to research or learn. Among those systems, reputation based method is a promising one which usually combines user reputation, bias, behavior model and relationship among different users to efficiently filter information and present the result satisfied the user’s demand. Recommendation is a strong incentive for users because better result can only be presented to the users of better reputation.

Our scheme is a modified version of PageRank and EngerTrust and takes users, papers, comments and private relationship into reputation computation. All the data above can be easily obtained in the platform by systematic means. This paper will examine how to integrate multiply clues of different qualities into a complete reputation metric and discuss the feasible reputation evaluation algorithm. The detail of the academic platform is described firstly in next section.

III. BASIC ACADEMIC RESEARCH PLATFORM

The platform includes an academic forum and P2P user network. This forum is thought exchange platform for the researchers to study literatures. In this forum, each literature is a basic item and other researchers post replies to feed back personal assessments. In fact, the researchers’ explicit feedbacks can serve as a kind of literature reputation, which can evaluate more accurate by combining impact factor of the source of the literature with the author's reputation, comments, comment reputation and so on.

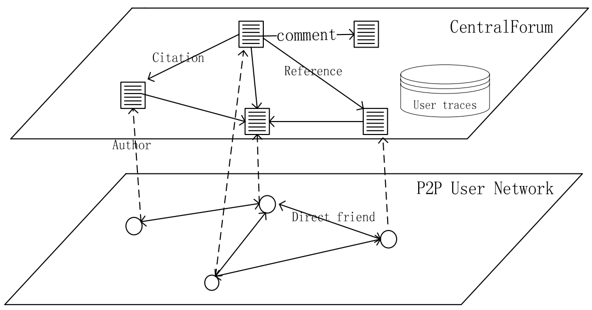

Fig. 1 Platform Structure

This forum is a public platform for the researchers to exchange their ideas. Forum database stores literatures that have been formal published in journals or conference proceedings. Then other users can search for them and get the detail information about them from the web page of the platform. Users can also post comments of personal opinions about articles or about the comments posted by other users. When a user posts a comment, he is enforced to mark the commented in the same time. Marks can be regard as the weight of recommendation.

A user can use the P2P tool to directly communicate to another user if the later is willing to interact with him. This is a kind of private exchange that can only be seen by participators. However, it is very important for reputation evaluation because the user can evaluate the partner not only by the explicit reputation, but also by the personal judgment about the partner through reading the articles the partner has published and interaction history with him. So the judgment is usually more pertinent than calculated reputation.

Reputation module is the core of the system which computes the reputation by interacting with multiple datasets and software modules and stores the results in the reputation database, as shown in Fig. 2.

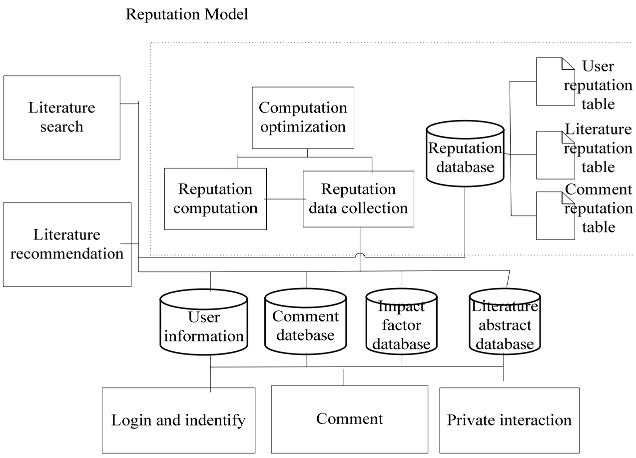

Fig. 2 Modules and structure of the academic platform

A search engine is employed to provide users with literature filter. The user needs to input key words and restrictive conditions firstly. The system works out literature list sorted by reputation then. When the user select one or more items to read, the system records the hits and the executor of each hit on background. Search and peer user selection use reputation data in reputation database. Search, comment and P2P communication apply network and user interaction module to transfer data and show the result on the screen.

In order to improve the accuracy of reputation evaluation and relieve the cheat and malicious behaviors, user identification divided into two classes: the authorized and the unauthorized. The user of the authorized is the author of a formal published paper and has passed the email check: Firstly, the checker sends an authentication message to the email address which is the corresponding address of the paper; if the checker receives a reply from the address later, he labels the user authorized user. Otherwise, the user is a unauthorized one.

IV. Reputation Computation Algorithm

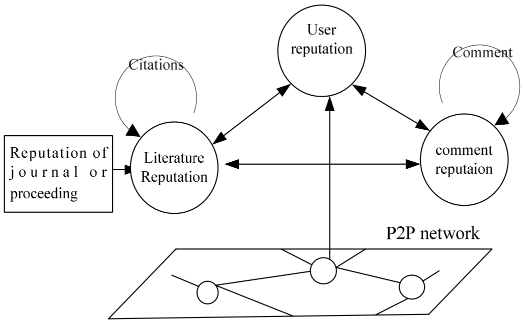

Fig.3 Reputation model

The literature, user or comment each has its own reputation, but each is related to others and its reputation is depended on the others' reputations, as shown in Fig.3, which is quite different to PageRank. The later has only one kind of element, web page. So the basic strategy is to compute local reputation firstly and to integrate them then. Iterate the process until each element reaches a

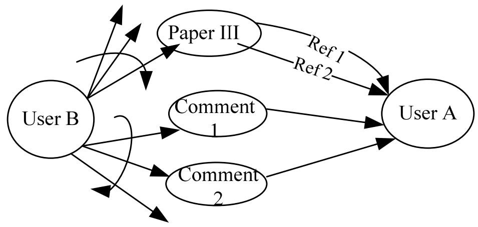

Fig.4 Construction of recommendation relation graph

stable status. P2P network is partly independent system, but it also contributes to reputation computing. Although users can gain their reputations only depending on relationships in P2P network, the reputations can also be introduced into reputations in the forum,.

The above model is too complex, so in order to effectively calculation user reputation, it needs some simplification. Referring to PageRank and EngerTrust reputation model, literature review or the citation can be regarded as recommendation relationship which weight is decided based on the presenter's reputation and the total number of recommendation, as shown in figure 4.

A. *Literature Reputation Model*

Literature reputation R_p shall consider at least three factors: authors, citations and comments. The formula (1) shows the reputation evaluation of literature i:

$$R_p(i)=\lambda_1 R_s(s(i))+\lambda_2 R_a(a(i)) +\lambda_3 R_c(i)+ (1-\lambda_1-\lambda 2_1-\lambda_3)R_r(i) \quad (1)$$

Here, d is a decay factor which is usually set to 0.85. R_s is the initial reputation of the journal or proceeding the article published on. Usually it is in proportion to its impact factor. If it has not impact factor, its initial reputation is 0.

R_a is the reputation of the first author, defined in formula (3). $R_c(i)$ is the score from comments about it, as show in formula (4).

R_r denotes citation value of the article.

$$R_r(i)= \sum_{j \in ref(i)} \frac{R_p(j)}{|ref(j)|} \quad (2)$$

$|ref(j)|$ donates the number of the citations of literature j.

B. *User Reputation Model*

User reputation reflects the user's academic authority and the academic value of his articles and reviews. User reputation has two sources: the forum and P2P network. There are two kind of user reputation: reputation of author and common reader. Fig.4 shows that the relation between two users is built on comments or articles. Therefore, we can calculate the two kinds of reputation by formula (3):

$$R_a(u)=\frac{\lambda_4\sum R_p(i)+(1-\lambda_4)\sum R_c(j)+R_a^p(u)}{2} \quad (3)$$

Here, i is a paper issued by the author u, j is a comment about u. $R_a^p(u)$ is the reputation gained from P2P network. We simply assume the reputation from forum is same important as that from P2P network.

C. *Comment Reputation Model*

Comment reputation of a comment k includes two parts: commentator reputation and the comments on this comment by other readers.

$$R_c(k)=\lambda_5\frac{R_a(a(k))Val(p(k),a(k))}{|comments(a(k))|} + (1-\lambda_5)\sum_{i\in sub(k)}\frac{R_c(i)Val(k,i)}{5|comments(i)|}) \quad (4)$$

$|comments(a(k))|$ denotes the total number the user $a(k)$ has posted, $Val(p(k),a(k))$ is the score $a(k)$ marks on $p(k)$ and $|comments(i)|$ is the total number of the comments posted by user i.

The comments form a tree structure that a father comment can be made up of a few child comments and each child may have a few children of him. The reputation formula is a recursive function: the end node in the comment tree is firstly evaluated by only R_a, his father then combines R_a and the reputations of child comments gained just now. The variables in the formulas above list in table 1.

D. *Reputation Computing in P2P Network*

In P2P user network, each user has several friends and all the friend relationships construct a friend network. Each edge in the network has a weight equal to the reputation of the partner, so PageRank likely algorithm can be used here. But different to PageRank, friend relationship is bidirectional, which needs to transform into two directed relationships with different directions, as shown in Fig.5. Assume that a is a friend of b, the reputation of a is denoted by $R_a^p(a)$. out(a) denotes the number of friends of a. So when the link between a and b is divided into two directed connects, the connect from a to b has a weight of $R_a^p(a)$ /out(a).

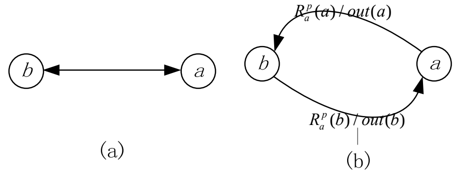

Fig.5 transformation of friend relationship

Reputation can be computed by an iterative algorithm. In the algorithm, the basic formula of u's reputation is:

$$R_a^p(u)=\sum_{v\in friend(u)}\frac{R_a^p(v)}{out(v)} \quad (5)$$

In the iteration, u is assigned an initial reputation equal to $R_a(u)$, the reputation of u computed by forum module.

The evaluation algorithm is detailed below:

$\forall u\in S \quad R_a^p(u)_0=R_a(u)$

//S is a user set where a user has at least one friend.

while($R_a^p(u)_i - R_a^p(u)_{i-1} > \varepsilon$){

for each $\forall u\in S$

$$R_a^p(u)_i=\sum_{v\in friend(u)}\frac{R_a^p(v)}{out(v)}$$

}

Because all nodes have at least one friend, they have at least a directed connect to other users. So, there is no dangling problem in the network.

E. *Dangling Problem*

Similar to the problem of dangling page in Pagerank, when a user hasn't any paper citing other papers and no comment, he is called dangling user. Simultaneously, the dangling paper is the paper without any citation. A comment always has an out link so it hasn't the dangling problem. Dangling entities can disturb the reputation evaluation of other entities.

Two methods can solve the problem to ensure the reputation computation convergent: (1) add a virtual link from the entity to all other available entities, the user can link to literatures and comments, the literature can link to literatures; (2) ignore all converse links pointed to the entity firstly in computation until he add a new link to another entity, for example, he posts a comment.

In fact, it is unimaginable that a paper without any citation is high in quality. So deleting the paper directly has little effect to the accuracy of reputation evaluation.

F. *Globe Reputation Algorithm and Convergence*

Evidently, reputation variables in (1)~(4) is not independent, so reputation computing is an iterative process and must be convergent. The adopted algorithm refers to the idea of recommendation networks. The evaluation algorithm is detailed in Fig. 6.

10 Procedure Evaluation
20 Begin
30 For each user u and literature i, $R_a(u)$=1 and $R_p(i)$=1
40 calculate initial reputation $R_p(i)$ by formula (1), (2) and (5)
50 store reputations of all entities t in $R(t)$
60 recalculate user reputation $R_a(u)$ with formula (3)
70 recalculate literature reputation $R_p(i)$ with formula (1)
80 recalculate comment reputation $R_c(k)$ with formula (4)
90 store reputations of all entities t in R'(t)
100 if $\max_{\forall t}\|R(t)-R'(t)\|>\varepsilon$ goto 50
110 end

Fig. 6 The iterative reputation evaluation algorithm

G. *Tradeoff Scheme*

The simulation below proves the convergence of the algorithm, but within large computing complexity. A tradeoff scheme proposes that the reputation computing steps, such as step 60, 70, 80, can only trigger when a new event happened. The event may be a submission of literature, comment or a register of user or periodical. When it happens, the corresponding formula will be called for one time but will not trigger the iterative process. It means that only the reputation of the entity the new entity directly links to changes and the change don't diffuse to other entities.

The reputation renew is local that mitigate greatly the computing cost. On the other hand, the tradeoff scheme cannot guarantee the accuracy of reputation that the speed of convergence to globe reputation is subject to events relative to this entity. To speed the convergence, the globe iterative algorithm may run periodically.

V. INITIAL REPUTATION

Formula(1) can be used to compute initial literature reputation, here $R_s(j)$ is the reputation of journal or proceeding j can be gained by literature databases. By now we choose most well-known databases include SCI journals, EI source, the list of Chinese core journals of PKU. The normalized computing is as (5):

$$R_s(j) = \begin{cases} 1\text{-}0.5^{factor(j)} & j \in \text{SCI} \\ 0.3 & j \in \text{EI} \\ 0.1 & j \in \text{core journal list of PUK} \end{cases} \quad (6)$$

Then, with the initial literature reputation above, initial user reputation can be computed by (3).

An authenticated user has an initial reputation corresponding to the reputation of journal where he has published a paper. The unauthenticated user is assigned a reputation of 0.01.

$$R_a(u) = \begin{cases} R_s(j) & u \text{ is authorized and } j \text{ is the journal} \\ 0.01 & u \text{ is unauthorized} \end{cases} \quad (7)$$

VI. SIMULATIONS AND PERFORMANCE ANALYSIS

Simulations abstract literature indexes randomly from SCI and EI database to construct the test base database, which is selected from 1995 to now, mainly from the field of computer science. The first authors of the papers in the base dataset are added into test user set. The dataset contains approximately 33, 500 articles, with 510, 000 citations. There are about 15, 000 authors. Two authors are considered identical if their full names match. There were a few citations from an article to another article published outside the paper set that were removed.

In the same while, the test produces a number of users who are not the authors, and only issue comments but not publish new literatures. The proportion of the users who are not authors to authors is 2. There are some comments randomly issued for each literature, includes direct comments upon the literature and indirect comments that are comments to other comments. The numbers of direct and indirect comments follow the Poisson distribution of $P(1)$. Commentators are chosen randomly from all users. Here, λ_1=0.2, λ_2=0.3, λ_3=0.2, λ_4=0.8, λ_5=0.5. Comment mark level is an integer selected from 0 to 5.

We first experiment to check the convergence of globe iterative algorithm. The experimental papers are selected randomly from base dataset. The number of documents in base dataset is called dataset scale.

At first, we specify a series of dataset scales from 400 to 33, 500 and the globe algorithm to calculate the reputation. The end condition is ε<0.001. The experiment with each dataset runs for 5 times, counts the entities not satisfied with end condition and records the calculation time. That can be used to evaluate the algorithm complexity. The result is shown in Fig.7 and table 1.

This experiment proves that the algorithm can converge to a determinate value within different base

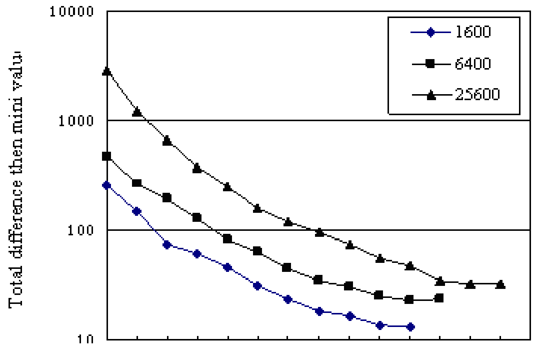

TABLE 1.

THE AVAILABILITY OF LOCAL ALGORITHM FOR REPUTATION EVALUATION

Dataset scale	Iteration rounds	time cost(s)
400	17.8	3.5
800	20.4	25.1
1600	22.8	64.2
3200	23.2	388
6400	25.6	1903
12800	27.2	7350
25600	28.8	34723

datasets in limited iterations that proves the availability of the algorithm. On the other hand, if the data scale increases, the calculation complexity increases dramatically. So the local algorithm is necessary.

The next experiment compares the globe algorithm to the local one. The local algorithm is called first. Once upon the update of the reputation x happened, the globe one will launch and get the exact value x* at that time. Initially, there are 100 items in the forum without any comments. In each round, the rate of new comments or new literatures added into the forum is denoted s. The comparison begins from 100 rounds.

$$deviation = \frac{|x - x^*|}{\max(x, x^*)}.$$

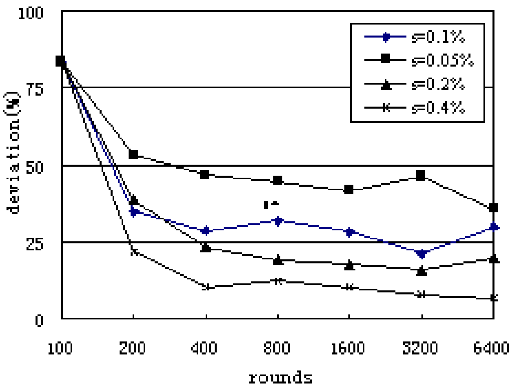

Fig. 8 The availability of local algorithm

Fig. 8 shows that after the test started, the accuracy of the local reputation evaluation kept on improving. s is the update speed of reputation in fact. When s arises, the evaluation precision increases. But the improvement cannot continue when the deviation is below a degree. That means the local algorithm cannot gain better performance if other methods aren't adopted. A feasible way is to collect all information and compute the reputations of all entities periodically, for example at weekend or after ten o'clock at night every Monday.

VI. Conclusion

We present a new reputation evaluation mechanism that can be used to evaluate reputation of literatures and researchers. The basic idea is that there are some relationships among literatures, its author, periodicals and readers. The relationship can be regard as recommendation for each other that is in some ways similar to hyperlink in Google. Therefore, we build a similar reputation model here and propose a PageRank-like algorithm. By simulations, we prove that the mechanism is feasible and convergent. We also discuss the strategies to apply the algorithm to gain better performance. A tradeoff scheme is proposed to replace the globe reputation by a local one. The outstanding advantage of local scheme is its simplicity and decentralization. Therefore, it is easy to deploy in a huge forum system.

There are some issues remained unsolved. One issue is if the reputation model and the evaluation method can exactly indicate the academic values of literatures or researchers. What is the standard? In despite of impact factor is widely accepted, but whether it is the most scientific one is in doubt. Another problem is how to decide the values of the parameters, such as $\lambda_1 \sim \lambda_5$. Here the values are specified subjectively. Lastly, because the comment is a key factor of reputation model, how to promote users to issue their reviews is the next work in the future.

References

[1] Alfarez Abdul-Rahman, Stephen Hailes. Supporting Trust in Virtual Communities. Proceedings Hawaii International Conference on System Sciences, 2000, 4 -7.

[2] Kaltenborn K-F, Kuhn K. The journal impact factor as a parameter for the evaluation of researchers and research, Rev Esp Enferm Dig. 2004, 96:460-476.

[3] L.Page, S.Brin, R.Motwani, and T.Winograd. The Pagerank Citation Ranking: Bringing Order to the Web, Technical report, Stanford Digital Library Technologies Project, 1998.

[4] Brumback RA. Worshiping False Idols: The Impact Factor Dilemma: Correcting the Record, Journal of Child Neurology, 2008, 23:1092-1094.

[5] Huth EJ. Authors, editors, policy makers, and the impact factor. Croatian Medical Journal, 2001;42(1):14-17.

[6] Sevinc A. Manipulating impact factor: an unethical issue or an Editor's choice? Swiss Med Wkly, 2004, 134(27-28), 410.

[7] Graczynski MR. Personal impact factor: the need for speed. Med Sci Monit, 2008, 14(10): ED1-ED2.

[8] Soualmia LF, Darmoni SJ, Le Duff F, Douyere M, Thelwall M. Web impact factor: a bibliometric criterion applied to medical informatics societies' web sites. Stud Health Technol Inform, 2002;90:178-183.

[9] Burke, R. Hybrid Recommender Systems: Survey and Experiments. User Modeling and User-Adapted Interaction,2002,12 (4),331-370.

[10] N. Ma, J. Guan, and Y. Zhao. Bringing pagerank to the citation analysis. Inf. Process. Manage. 2008, 44(2):800-810.

[11] Srivastava, J., Cooley, R., Deshpande, M,Tan, P. Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data. SIGKDD Explorations ,2000, 1(2),12-23.

[12] S.D.Kamvar, M.T.Schlosser, and H.G.-Mollina. The Eigen Trust Algorithm for Reputation Management in P2P Networks, The Twelfth International World Wide Web Conference, 2003.

[13] Min Peng, ZhengQuan Xu, ShaoMing Pan, Rui Li, Tengyue Mao. AgentTMS: A MAS Trust Model based on Agent Social Relationship, Journal of Computers, 2012, 7(6), 1535-1542.

[14] Hui Chen. The Impact Mechanism of Consumer-generated Comments of Shopping Sites on Consumer Trust, Journal of Computers, 2011, 6(8), 1677-1682.

[15] G Adomavicius , A Tuzhilin1. Toward the next generation of recommender systems : A survey of the state of the art and possible extensions, IEEE Trans on Knowledge and Data Engineering, 17(6),734 - 749,2005.

Kun Yu Born in Huaian, Jiangsu, China, 1972. He Received the Ph.D. degree in computer science and technology from southeast University, China in 2010. Currently, he is a teacher at Huiyin Institute of Technology, China. His research interests include computer network and protocol analysis.

Jianhong Chen Born in Huaian, Jiangsu, China, 1972. He received his Ph.D. degree in Computer Science and Technology from Shanghai Jiao Tong University in 2011. Currently, he is a teacher at Huiyin Institute of Technology, China. His research interests include public key cryptosystem and network security.