The Searching Ranking Model Based on the Sharing and Recommending Mechanism of Social Network

Fei Hongxiao, Mo Tianchi, Wang Yang, Wu Zequan, Liu Yihuan
School of Software
Central South University
Changsha, China
Email: {hxfei, motianchi, wangyangcsusoft, wuzequancsusoft, 3901130215}@csu.edu.cn

***Abstract*—The combination of social network and search engine is the trend of Internet in coming years. Through introducing the widely utilized sharing and recommending mechanism in social network to search engine, this paper proposes a new searching ranking model. This model judges the quality of web pages and decide what extent do they meet users' personalized need through analyzing the records of users' social circle's behavior of sharing and recommending. Then, it can make search engine provide users with personalized results sequences. Both the experiment and the theoretical analysis show the proposed model can automatically help users to select the high quality search results, and provide users with better personalized service.**

Keywords—Search Engine; Personalized Search; Ranking Model; Social Network; Sharing and Recommending Mechanism

I. INTRODUCTION

Search engine (SE) and social network are the two most active kinds of application systems in present Internet. Google and Baidu are the most popular SEs, and both of them are belong to the second generation full-text search engine. The advantages of this kind of SE include fast speed, high recall ratio, timely updating, and total automation. However, with the sharp expansion of the online information quantity, some obvious defects of this kind of SE have been gradually exposed. Their precision is not very high, and many unrelated webpages, or the webpages that not accurately satisfied users' need are also contained in the result list, so, users have to spend lots of time in selecting the pages they really want. These SEs are also lack of proactivity, and not capable of providing precise personalized services [1]. Simultaneously, in the era of Web 2.0, social network, like Facebook and Twitter, has been rapidly ascending in recent years. Due to users spend much more scattered time in communicating on the social application, the social webs can accumulate massive data closely related to users' characteristic. It makes social networks able to supply more real-time and more personalized services to users. As a matter of fact, social network is becoming the new major source of information after the time of SE appearance [2]. In the future, it may overstep SE and become the most important traffic entrance of other websites [3]. Nevertheless, there is randomness and uncertainty when people obtain knowledge online, and their social relations have boundaries, that cause the information given by the social web cannot cover whole searching requirement of single user.

SE and social network both have pros and cons. The combination of them is the trend of cyber application. Literature [4] points out that the proactivity, personality and sociality of social application can make up the defects of modern SE. And SE's ability to gather resource from whole Internet makes it capable of providing much more information than social webs. Therefore, in next decade, SE may comprehensively combine with social network. The new mixed application will possibly become the most critical online system in daily life.

So, in this paper, to make SE become more intelligent, accurate and personalized, we introduce the common sharing and recommending mechanism in developed social nets, and try to transplant it into search systems. A new ranking model named Sharing & Recommending Result Rank (SRRR) is proposed in this purpose. SRRR is designed to help users select search results. It can return different webpage lists to different users, which is the real personalized searching based on social information.

The rest of the paper is organized as follows. Section 2 reviews the related work about the combination of social web and SE. Section 3 introduces the proposed SRRR model. Theoretical, empirical and experimental analysis is presented in Section 4. And finally, the conclusion and the future work are given in Section 5.

II. RELATED WORK

Actually, industry does not have consensus on the definition of "social search". Literature [3, 5, 6] deem social search to be the search function in the social application, which means when users submit the keywords, systems return the related contents (like tweets, photos, blogs and so on) with order. Meanwhile, other scholars regard the core of social search as the full-text S, which is faced with the whole Internet. In their opinion, social search means using the intelligent data of social web to comprehensively optimize the second generation SE, rather than searching the social website itself [1, 2].

Academic research about the combination of SE and social application began with the deep exploration of bookmarking sites [7, 8]. Bao et.al put forward the social similarity rank model (SSR) and the social page rank model (SPR) [7]. Both SSR and SPR are based on the del.icio.us, and the hypothesis that people search pages through annotations. SSR is an iterative algorithm to identify the similarity of annotations through checking the people who use them, while SPR extends the page rank algorithm to the network consist in pages and annotations. Chia-Hao Lo et.al pose the expert vote rank (EVR) and recommendation page rank (RPR) [8]. EVR is based on users' prestige in del.icio.us, and RPR uses the weighted directional recommendation graph to do the ranking. After above approaches, some algorithms and models have been designed for the search function in social platforms. Xing et.al propose that RDF data can be used to solve the heterogeneity problem in social network searching [5], but they do not design an effective rank model. And Xu et.al put forward a ranking model involving the page publishing time, the relevance of content and key word, the page publishers' prestige and the relation between publisher and searcher [9]. However, this model is confined in Twitter and only covers the data of two months. As for the optimization of full-text search engine with some social mechanisms, there is few study achievements so far. And some companies have developed productions, but their effects need to be checked by time. Facebook gets a patent called "curated search", while they also cooperating with Bing, the SE developed by Microsoft [3]. Google also designs Google+, hoping to build their own social platform. In academic area, Toyokazu Akiyama et.al also proposed a method through extending page rank. They apply the page rank to a new class of nets with two kinds of vertices, pages and people who are reading these pages right now. It is very creative, but it is too hard to implement.

As far as we can see, academia is now focusing much more on the site search of certain social platform, and the improvement of page rank with shallow social information. However, until today, there is no effective approach to incorporate the sociality into the full-text search engine on a high level. As mentioned above, the functional limitation makes social web cannot totally replace full-text search engine. Frequent searching behaviors need to be supported by the data from whole World Wide Web. For this purpose, the design of SRRR aims at using the superiorities of social network to prove the effect of SE. The application environment will not be confined in certain social platform.

III. Using Social Sharing & Recommending Mechanism to Optimize Search Ranking

Virtually, as literature [3] pointing out that what threatens the SE is not the social network itself, but the intelligent data of the social web. Users spend lots of time on interacting with social applications and other users, so these applications can accumulate massive data related to users' information. Therefore, social applications get better know about their user. Without users' active expression, we can use the data to build interest vectors [1]. Generally, social nets provide the completed mechanism of sharing and recommending, large quantity of information can be diffused rapidly along the links of web. If we can utilize the features hiding in the actions of sharing and recommending guiding search ranking, then the SE will become more personalized due to the personality of everyone's social circle.

A. Relevant Conceptions of SRRR

In existing social application, the conceptions related to SRRR include social relation, social distance, and sharing & recommending mechanism.

1) Social Relation: User u and v having social relation on a social platform means an explicit record of the edge between 2 vertices in the database. There are 2 kinds of relation, unidirectional and bidirectional. Unidirectional relation means if there is a link from u to v, the link from v to u is not necessary. In Twitter and Sina micro blog, "Follow" is a kind of unidirectional relation. We denote unidirectional relation as $u \rightarrow v$. Meanwhile, in Facebook and Tencent QQ, the social relation is bidirectional, that means if there is a link from u to v, the link from v to u also must exits. It is denoted as $u \longleftrightarrow v$. The proposed model only takes unidirectional relation into consideration, because $u \longleftrightarrow v$ is equal to $u \rightarrow v$ && $u \leftarrow v$. The conception of social relation maps the social network as a directed graph.

2) Social Distance: In existing social nets, information is one-step-visible [5], which means u can read contents posted by v iff $u \rightarrow v$. Social distance is defined for Information transmission. If $u \rightarrow v$, the social distance is 1, denoted as $dis(v, u)=1$. If $dis(v, u)=d$, $w \rightarrow u$, w not$\rightarrow v$, and $\forall w \rightarrow x \&\& x \neq u$, $dis(v, x) \geq dis(v, u)$, then $dis(v, w)=d+1$. Due to the huge number of users in social web, ordinary algorithms are incapable of finishing the calculation in short time. The method proposed by literature [10] can be taken to curtail the run time.

3) Sharing & Recommending Mechanism: Most of the social applications provide the service of sharing and recommending. In different positions of different webpages, we can see the sharing button. And in social applications themselves, users also can use "Forward" function to share the contents they like. If $u \rightarrow v$, then, when v has retransmitted something u will be noticed and can read the content v shares. But to users who do not follow v, they will not directly know v's behavior.

Recommending, also known as "Like", is a kind of behavior visible to whole site. The application will record users' positive evaluation to certain content. When other users request a set of data, the content with more recommendation is nearer to the top of result list.

B. Process of SRRR

Nowadays, the second generation full-text SE works with following procedure: collecting webpages, analyzing webpages, building the inverted index, and handling the queries [11]. The proposed model will influence all of these steps. The thought of SRRR comes from users' daily behavior of sharing and recommending. Normally, only the pages / resources with high quality and exactly satisfying users' requirement can get users' "Forward" and "Like". And the behavioral basis of sharing and recommending is user and his social circle having analogous interests. Therefore, SRRR is based on following hypothesizes:

- Hypothesis 1: The shorter the social distance between two users is, the more similar the requirement they share [1]. Namely, user and his social circle have similar interests.
- Hypothesis 2: Page's quality is proportional to the times it has been shared. Simultaneously, if the user who shared the page has high prestige in social web, this certain sharing behavior should have higher importance. The more a content has been retransmitted, the nearer it is to the top of the result list in searching.
- Hypothesis 3: Page's quality is proportional to the times it has been recommended. Meanwhile, to certain webpage, if its recommenders have shorter social distance to the searcher, and the number of them is larger, the page's position should be nearer to the top of return list.

Through calculating the Share Ranking Factor (SRF) and the Recommending Ranking Factor (RRF), SRRR combines the attitude to webpages of searcher's social circle with other elements of search ranking, and helps users to make decision when selecting search result.

1) SRF Calculation: Before calculating SRF, the prestige of users in social network needs to be identified. We call it Prestige Factor (PF). Page Rank algorithm [9, 12] is taken to solve this problem. The primary idea of this algorithm is: the more users follow *u*, the larger *u*'s PF is. Meanwhile, if *u* has large PF and $u \rightarrow v$, *v* also has large PF. Equation (1) is utilized to calculate PF:

$$pf(u) = (1-d) + d\sum_{v\in f_u}\frac{pf(v)}{N_v} \tag{1}$$

In Equation (1), $pf(u)$ denote *u*'s PF, $d \in (0, 1)$ is attenuation factor, usually, $d=0.85$. And f_u is a set of users, $\forall v \in f, v \rightarrow u$. N_v is the number of users followed by *v*. Application system can measure users' PF when its load is light, like 12.p.m~4.a.m every day. This is a preprocessing step, so it does not cut down the system performance. When user *u* submits the search request with keyword *k*, system firstly retrieves out *pages*, the set of webpages whose keyword is *k* (existing SE has this step, too). Then, Equation (2) is taken to calculate SRF of every $p \in$ *pages* for *u*:

$$srf(u,k,p) = \sum_{v\in(bf_u\cap Share(k,p))}\frac{pf(v)}{N_u} \tag{2}$$

In Equation (2), bf_u is a user set. $\forall v \in bf_u, u \rightarrow v$. *Share* (*k*, *p*) is also a user set. It has two kinds of elements. One includes the users who directly share p into social net without using the SE, and after analysis the system reckon k as the keyword of *p*. The other contains the users who share *p* when they searching keyword k through SE. Nu is the number of users followed by *u*.

2) RRF Calculation: As the definition in Section 3.1 (3), recommending, or "like" is visible to full site. User *u* can recommend the webpage *p* he likes to the whole social web. According to Hypothesis 1 and Hypothesis 3, proposed model computes RRF with following way:

Build a pattern M, which is a triplet (*u*, *k*, *url*). It means when *u* searching keyword *k*, he recommended one of the results whose web address is *url*. Or u directly recommended the webpage whose address is *url* to the system, and system regards *k* as its keyword. When user *u* submits the search request with keyword *k*, same as SRF calculation, system figures out pages, then, uses Equation to measure RRF of every $p \in$ *pages* for *u*:

$$rrf(u,k,p) = \sum_{v\in Rec(u,k,p)}\frac{threshold - dis(u,v) + 1}{num(u, dis(u,v))} \tag{3}$$

$Rec(u, k, p)$ is a user set. $\forall v \in Rec(u,k,p)$, there must be a record in M values (*v*, *k*, *url* of *p*), and $dis(v, u) \leqslant threshold$. In Equation (3), threshold is a constant number, which denotes the upper bound of social distance. According to Six Degrees of Separation [6], we set $threshold=6$. Users with longer social distance will be ignored. Meanwhile, suppose $dis(v, u)=d$, $num(u, dis(v, u))$ is the number of users whose social distance to *u* is also *d*.

3) Search Ranking Based on SRF and RRF: After SRF and RRF are obtained, combine them with existing ranking algorithm by weighting. Final step to get ranking value needs the following equation:

$$srrr(u,k,p) = \left(\sum_{i=1}^{n} c_i f_i\right) + c_s srf(u,k,p) + c_r rrf(u,k,p) \tag{4}$$

In Equation (4), $srrr(u, k, p)$ is the ranking value of webpage *p* when user *u* search keyword *k*. The webpage with bigger SRRR value can get closer position to the top of the result list. Meanwhile, $c_1, \ldots, c_n$, c_s and c_r are weighting factors, $f_1, \ldots, f_n$ is the n ranking factors (like page rank, timing factor, bounce rate and so on) already utilized by current SEs. To any f_i ($1 \leqslant i \leqslant n$), if pages with bigger f_i can be nearer to the top, then, c_i is positive, otherwise, c_i is negative. Because both SRF and RRF are the larger the better, c_s and c_r are positive. In this case, SRF and RRF are capable of affecting the ranking result as importantly as other factors like page rank value. If searching system needs to be influenced more by social information, the value of c_s and c_r should be amplified. Different users have different social circle, and the attitude and emotion tendency of user's social circle can help user to make personalized decision. So, it can be concluded that with sharing & recommending mechanism, SRRR is able to make the ranking of full-text SE more in line with users' personalized requirement.

IV. Analysis of SRRR

The rationality of SRRR will be analyzed in theoretical and empirical aspect, and then, a simulation experiment will be presented.

A. Theoretical and Empirical Analysis

What has to be mentioned is most of the existing social search approaches are focused on the adaptation of page rank.

It is not very reasonable. Firstly, page rank algorithm was designed for static Internet in late 1990's. Today's Internet is dynamic. Page rank evaluates webpages through links between them, which only based on the assessment of developers, not scanners or readers. To some web sites that are not good at link optimization, even though they may have content with high quality, their page rank value may be very low. This problem also bothers the new born websites. Meanwhile, some sites with high SEO capability might cheat SE by using spam links, that makes users get many irrelevant and inaccurate information when doing their search [6]. No matter how to change the page rank algorithm, above defects will not be fixed as long as the information construction of Internet is based on hyperlink. Under this frame, page rank value is an inherent property to the webpages or sites.

Secondly, although page rank is not perfect, it is still a very effective ranking method. Most of SEs takes it as a core ranking factor all over the world. Modification of page rank may change the kernel systems of SEs, which requires huge workload. In addition, developing a system with both search service and social service may be not a good idea. It is hard for new application to gather users in the area of social web and SE today. And all of these thoughts are contradictive to the Open Closed Principle of software engineering. Finally, page rank is only one of many ranking factors, its weight is not very high. Take Google as an example, when sorting the search results, there are dozens of factors influencing the final order, including page rank value, page view value, bounce rate, response rate and so on. Just combining the social information with page rank cannot make the social data affect ranking comprehensively.

Additionally, considering the personalized searching, the page rank value of website is not very meaningful. As mentioned, page rank value is an inherent property of webpages. It is based on the hyperlinks, and relatively static. It does not change with different users and keywords. In other words, page rank is not a personalized algorithm. So, when the personalized SE is being designed, we need find a new approach to utilize the social data independent to the page rank model.

The proposed model, SRRR, does not have the defects mentioned above. Through analyzing Equation (2), it can be found that to different users, one webpage has different SRF. The number of user u's friends does not have effect on SRF calculation and only the behavior of users directly followed by u will be involved into the calculation. At the same time, SRF and the prestige of sharers have positive correlation. So, Equation (2) is in line with the Hypothesis 1 and Hypothesis 2, which make SRF represent users' personalized requirement objectively. And Equation (3) takes the number of recommenders and social distance into consideration, which meet the requirement of Hypothesis 1 and 3, reflecting users' need through analyzing the recommending behavior in social circle. Like SRF, the same page has unequal RRF according to various people. More importantly, Equation (4) defines that SRRR model is a complementary optimization to existing SE. When using SRRR, the modification of internal structure for existing systems is not necessary. What is necessary is only adding two factors with proper weight. So, SRRR has high usability, acceptable to the software engineering. Because we consider the social factors as independent ones, when the social network is not mature (if the SE enterprises choose building social web by themselves), the weight of SRF and RRF can be lowed down to make sure the old ranking result from existing models will not be disturbed. In fact, if there is no "Forward" or "Like" behavior, the output of SRRR will be same with the existing algorithm. Further more, with the development of social net, the weight of SRF and RRF can be amplified, then, social behavior can affect search ranking globally. With SRRR, new webpages, and the sites with good content and poor SEO can get better rank rapidly, and the search optimization with social information and personalized search will be truly implemented.

It is worth mentioning that when exploring the way to combine social network with SE, what has been found very hard is how to protect users' privacy. On this aspect, SRRR only uses the public behavior of people instead of gathering and analyzing users' sensitive information deeply. Meanwhile, when designing SRRR, we fully consider about the different visibility of sharing and recommending. We use information from full site to calculate RRF, however, when calculating SRF, only users' friends will be involved. So, SRRR will not cause privacy problems.

B. Experimental Verification

Unfortunately, there is no developed platform with both search function and social service. So, it is impossible to get real data for experiment. To finish the experimental verification for SRRR, we build a tiny dataset to stimulate the situation. It is supposed there is a social-SE. Some users in this system form a network as Fig 1 shows.

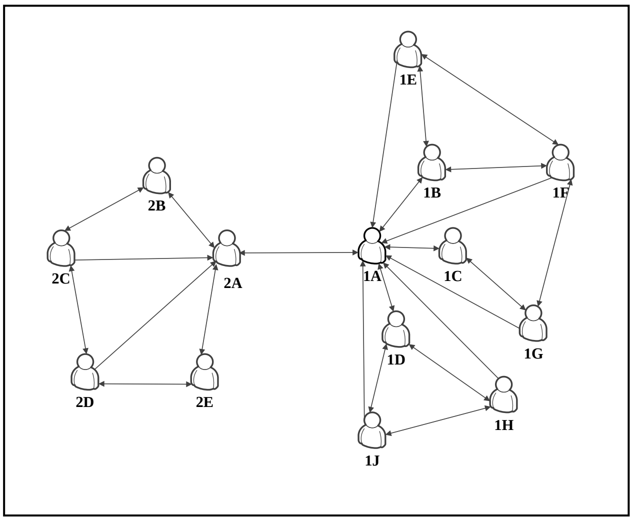

Fig. 1. Users' social network.

In this web, users with ID starting by 1 is a software develop team (referred as Circle 1), and users with ID from 2A to 2E are some fruit peasants (referred as Circle 2). Peasant might ask the team to produce an APP to promote their apple, so, there might be links between two circles (2A←→1A). When these 14 people search word "apple" by Google or Bing, they will get the results with same following order:

$$R_{original}=\{R_{f1},R_{p1},R_{p2},R_{p3},R_{f2},R_{p4},R_{f3},R_{f4},R_{f5},R_{f6}\}$$

In $R_{original}$, left members have the position nearer to the top than right ones on the output page of SE. Briefly, these results (webpages) are divided into two categories. R_{fi} pages are relevant to apple cultivation, while R_{pi} pages are related to Apple Inc. (iPhone, iPad, et.al). Index i means this page get the i^{th} position in its category.

Obviously, Circle 1 wants R_{pi} more, while Circle 2 is more concerned about P_f pages. Now, in social-SE, suppose users will share or recommend the result pages they like or feel helpful, like TABLE I shows. We will exclude user 1G and 2C.

TABLE I. RECORDS OF SHARING AND COMMENDING

User ID	Pages	
	Sharing pages	***Commending pages***
1A	R_{p1} R_{p3}	R_{p1} R_{p3} R_{f4}
1B	R_{p2} R_{p3}	R_{p3}
1C	R_{p1}	R_{p1} R_{p3}
1D	R_{p3} R_{p4}	R_{p3} R_{p4} R_{f6} R_{f1}
1E	R_{p2} R_{f6} R_{p4} R_{p1}	R_{p2} R_{p3}
1F	R_{p3}	R_{p3}
1H	R_{p1} R_{p3}	R_{p1}
1J	R_{p4} R_{p3} R_{p1}	R_{p3} R_{p1}
2A	R_{f4}	R_{f4} R_{f1} R_{f2}
2B	R_{f4} R_{f5}	R_{f4} R_{f5} R_{p2}
2D	R_{f4}	R_{f4} R_{f6} R_{p3}
2E	R_{f1} R_{p3}	R_{f1} R_{f3}

From TABLE I, it can be concluded that R_{p3} is popular in Circle 1, while Circle 2 prefers R_{f4}. Some noise is set in the dataset, like that 1E shared R_{f6} and 2D recommended R_{p3}. Briefly, the Equation (4) will be simplified as Equation (5):

$$srr(u,k,p) = c_1 f_1 + c_s srf(u,k,p) + c_r rrf(u,k,p) \quad (5)$$

In Equation (5), $f_1=11-i$, i means page p is on the i^{th} position in $R_{original}$. Here, f_1 is a comprehensive factor that covers all ranking factor in existing SE. Weight c_1 is given a small value to eliminate the error caused by the simplification. We use Java to implement SRRR, and set $d=0.85$ for Equation (1), $threshold=6$ for Equation (3) and $c_1=0.1$, $c_s=0.7$, $c_r=0.2$.

When taking 1G and 2C into consideration, SRRR values of different webpages to different users for searching "apple" are shown in TABLE II.

TABLE II. SRRR VALUE

Webpage	User ID				
	1A	***1B***	***1C***	***1D***	***1E***
R_{f1}	1.7429	1.7000	1.6600	2.8933	1.7000
R_{f2}	0.9000	0.8500	0.8500	0.9333	0.8500
R_{f3}	0.5429	0.6000	0.5600	0.5600	0.6000
R_{f4}	3.1422	1.7500	2.2454	1.7933	1.7500
R_{f5}	0.3429	0.4000	0.3600	0.3600	0.4000
R_{f6}	1.2000	1.1685	0.9500	2.100	0.9500
R_{p1}	3.0514	2.4726	4.0116	3.1794	2.2542
R_{p2}	1.4006	1.8685	1.7200	1.2800	2.8794
R_{p3}	4.3173	5.0861	5.4616	5.0928	5.3154
R_{p4}	0.9739	0.9685	0.7500	2.1210	0.7500

Webpage	User ID				
	1F	***1G***	***1H***	***1J***	***2A***
R_{f1}	1.8667	1.7000	1.8933	1.8933	3.2229
R_{f2}	0.9333	0.8500	0.9333	0.9333	2.0000
R_{f3}	0.6000	0.6000	0.5600	0.5600	0.8000
R_{f4}	1.901	1.7500	1.7933	1.7933	3.1229
R_{f5}	0.4000	0.4000	0.3600	0.3600	0.8229
R_{f6}	1.1972	0.9500	1.1000	1.1000	0.7000
R_{p1}	2.4930	2.6251	3.9585	3.9585	2.1242
R_{p2}	1.9359	2.6500	1.2800	1.2800	1.5200
R_{p3}	5.0345	5.0861	4.1036	5.1036	3.0671
R_{p4}	0.9972	0.7500	1.3528	1.1318	0.7000

Webpage	User ID			
	2B	***2C***	***2D***	***2E***
R_{f1}	2.5396	2.1667	2.516	3.2667
R_{f2}	1.2000	1.0000	1.000	1.2000
R_{f3}	0.7333	0.9000	0.8000	1.8000
R_{f4}	4.3050	4.1151	3.9922	3.2383
R_{f5}	1.6000	0.8229	0.7000	0.5333
R_{f6}	1.6396	2.1667	2.3930	1.3000
R_{p1}	1.7400	1.9067	1.9067	1.7400
R_{p2}	2.4400	1.4400	1.5400	1.3733
R_{p3}	2.6467	2.8800	4.1029	2.9133
R_{p4}	0.7667	0.7667	0.7667	0.7667

According to Equation (4) or (5), SRRR value is a function of user, keyword and webpage, and in TABLE II, it can be found to different users, one page has different SRRR value when the keyword is "apple". So, it proves that SRRR is a personalized rank model.

If the results are arrayed by SRRR value in descending order, the output lists of user 1G and 2C will like TABLE III and IV present:

TABLE III. SEARCH RESULT LIST OF USER 1G

NO.	Search result of User G1	
	Webpage	*SRRR value*
1	R_{p3}	5.0861
2	R_{p2}	2.6500
3	R_{p1}	2.6251
4	R_{f4}	1.7500
5	R_{f1}	1.7000
6	R_{f6}	0.9500
7	R_{f2}	0.8500
8	R_{p4}	0.7500
9	R_{f3}	0.6000
10	R_{f5}	0.4000

TABLE IV. SEARCH RESULT LIST OF USER 2C

NO.	Search result of User 2C	
	Webpage	*SRRR value*
1	R_{f4}	4.1151
2	R_{p3}	2.8800
3	R_{f1}	2.1667
4	R_{f6}	2.1667
5	R_{p1}	1.9067
6	R_{p2}	1.4400
7	R_{f2}	1.0000
8	R_{f3}	0.9000
9	R_{f5}	0.8229
10	R_{p4}	0.7667

Because 1G belong to Circle 1, so, as TABLE III shows, R_{p3} and R_{p2}, the webpages with much sharing and recommendation in this circle, are put on the top of 1G's list. As well as R_{f4} and R_{f1} in 2C's list. Namely, with SRRR, SE can choose the results through analyzing behaviors of users' social circle, find the good pages, and put them on the head of the list. In summary, compared with existing SE, the simulating system with SRRR is more personalized, meanwhile, it is capable to save users' time to find the best pages they really want.

CONCLUSIONS AND FUTURE WORK

This paper briefly introduces the research status of the combination of social network and search engine, and analyzes the advantages and deficiencies of existing study. Then, we propose a new model called Sharing & Recommending Result Rank (SRRR). Through mining the behavior of users' social circle, especially the sharing and recommending action, SRRR can gain the evaluation of users' social circle to the content, which make it able to help people choose the search result. To different users, SRRR can provide different search result ranking. Theoretical, empirical and experimental analysis is in line with the expectation. So, it is proved that SRRR is an effective approach with proactivity, personality and sociality, just as Literature [4] anticipated.

Like we mentioned in 4.1, it is very hard to build a new and fully developed social network in existing search engine. And until now, an application system with both full-text search function and social service has not been developed. So, in future research, we hope find a way to combine existing social web and SE, like Google and Twitter, or Baidu and Sina Micro Blog. The data from developed social net can make SE more intelligent, while the full-text SE can provide much more information to the social web. The approach to integrating their capability of allocating network flow will be explored, and SRRR might be the starting point.

Future work also include adapting and ascertaining the best range of SRRR parameters, analyzing users' interests, improving the effect of SRF and RRF calculation, and taking the specialization of users (expert users) into consideration. Further more, social search engine is based on the hypothesis that users will actively interact with the system. So, how to encourage users sharing and recommending contents, and how to prevent spam behavior (like malicious "Forward" or "like") is also need to be researched.

ACKNOWLEDGMENT

The research is supported by "National Natural Science Foundation of China" (No.61202095, NO.61073186) and "National Undergraduates' Innovation and Entrepreneurship Training Program of China" (NO. 201310533018).

REFERENCES

[1] CAO Jiandong, TANG Yang, LOU Binbin.Social search engine research[C]. 2010 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT), 2010, 3(7): 308~309 .

[2] AKIYAMA T, KAWAI Y, MATSUI Y. A proposal for social search system design[C]. The 11th IEEE/IPSJ International Symposium on Applications and the Internet (SAINT), 2011: 110 – 117.

[3] Li Fei. The Game between Search Engines and Social Networks [J]. Communications of the CCF, 2011, 7(9): 54~57.

[4] ELOWITZ B. Search and social: how the two will soon become one[EB/OL]. [2012-06-17] http://techcrunch.com/2012/06/17/search-and-social-how-two-will-soon-become-one/.

[5] Xing Chao, Zhang Wenbo, Zhang Xiang et.al. SNSearch: A Search Engine for Social Network Search [EB/OL]. [2013-01-05] http://www.paper.edu.cn/releasepaper/content/201301-172.

[6] Li Jing. The Development Trend of Next Generation Search Engine: Social Search [J]. Science of Information, 2010(13): 14.

[7] BAO Shenghua, WU Xiaoyuan, FEI Ben. Optimizing web search using social annotations[C]. Proceeding WWW '07 Proceedings of the 16th international conference on World Wide Web, 2007: 501-510

[8] LO Chia-Hao, PENG Wen-Chih, CHIANG Meng-Fen. Ranking web pages from user perspectives of social bookmarking sites[C].2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Workshops WI-IAT Workshops 2008, (Volume:1): 155~161

[9] Xu Jie, Kang Muning, Dong Guyin. The Research of Ranking Algorithm for Real-time Search Engine Based on Social Networks [J]. Science Technology and Engineering, 2011, 11(28): 6879~6882

[10] Shi Lingfeng. Research and Application of Graph Query Algorithm Based on Friends Relationship in Social Network [D]. Nanjing: Nanjing University of Science & Technology, 2012

[11] Pan Xuefeng, Hua Guichun, Liang Bing. Stepping into Search Engine (V2) [M].Beijing: Publishing House of Electronics Industry, 2011

[12] PAGE L, BRIN S, MOTWANI R, et al. The Page Rank Citation Ranking: Bringing Order to the Web[EB/OL].[1998-01-29] http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf

[13] VU T, BAID A. Ask, Don't Search: A Social Help Engine for Online Social Network Mobile Users[J]. Sarnoff Symposium (SARNOFF), 2012 35th IEEE, 2012(35): 1~5

[14] Li Deyi. Cloud Computing Technology Development Report 2013 [M]. Beijing : Science Press, 2012