Temporal PageRank on Social Networks

Weishu Hu[1(✉)], Haitao Zou[2], and Zhiguo Gong[3]

[1] School of Science, Jiangnan University, No. 1800 Lihu Avenue, Wuxi, Jiangsu,
People's Republic of China
huweiss@outlook.com
[2] School of Computer Science and Engineering, Jiangsu University of Science
and Technology, No. 2 Mengxi Road, Zhenjiang, Jiangsu, People's Republic of China
nkroben@outlook.com
[3] Department of Computer and Information Science, University of Macau Avenida
da Universidade, Taipa, Macau, China
fstzgg@umac.mo

Abstract. Social network has been a widely accepted way for people to communicate and interact online. However, few of existing works studied temporal dimension in assessing the authority of nodes on social networks. In this paper, a novel Temporal PageRank (T-PR) algorithm is proposed for analyzing the authority of nodes. Three temporal factors are adopted to personalize PageRank, which favors the nodes that are more important to people. They are Built-up Time-length Factor (BTF), Frequency Factor (FF), and Similarity Factor (SF). The experiments on a real data set demonstrate T-PR algorithm provides the best ranking results over recent competitor methods.

Keywords: Temporal ranking · PageRank · Link analysis · Search engine · Time-weighted ranking

1 Introduction

Over the past decade, the social networks have evolved from being an information source to a center of the world for commercial and social role. Web search engines are important to help users to find the most useful resources for their specific interest, which is to bring the most relevant web pages to the top ranked list for a given query. Most search engines include a ranking algorithm that computes a page's authority based on either the link structures of the web, e.g., PageRank [25] and HITS [17], or mining of users' web histories, e.g., BrowseRank [19], Traffic-weighted Ranking [22], and BookRank [9]. All of these algorithms produce rankings for different uses. But these algorithms may be biased against more recent pages [13, 24], while such pages have less time to accumulate in-links to contribute to their link-based ranking, and less chance to be involved into a web history.

However, an important factor that is not considered by these algorithms is temporal information which is critical to user's interest in other users. Social

networks are drastically different from traditional web, which are dynamic interaction environments. Quality users in the past may not be quality users now or in the future. In this paper, we study search from the temporal dimension, which is important due to the following reasons:

– Users are often interested in the recent or active user. Except for the users are celebrities or stars, most users on the network change constantly. New friends are added; ideally, outdated friends are deleted. However, in practice many outdated links are not deleted. This fact prevents the ranking algorithms from retrieving the update results.
– Existing Web page evaluation algorithms basically favor pages that have many in-links. Thus, older users are favored because they tend to accumulate more in-links due to longer existence. In contrast, new users that are high quality will not be ranked high.

We believe that dealing with the problems related to the temporal dimension is of great importance to future developments of rank technology in social networks. In this paper, we investigate the value of incorporating temporal aspects into ranking of users. Three temporal factors are considered: Built-up Timelength Factor (BTF), Frequency Factor (FF), and Similarity Factor (SF), which capture the intuitive notion that a user with recent updates, special time occurrence, or trend in revision is potentially more important to users. The hypothesizes of this paper are (1) the longer interval between registration time of user and creation time of link is more trustworthy; (2) trustworthy users are often active and regularly adding/deleting links, and seldom perform majority operations at once; (3) recent link is more important than old link.

Trust network is a directed graph, which has explicit links to express one node trusts/distrusts other nodes. In a trust network, there exists two important times (the registration time of node and the creation time of trust link). Individual users are represented by nodes with registration time, having the relationship "User X trusts User Y on Time t" resulting in an edge directed from User X's node to User Y's on Time t. Everything has its cause; there is no absolutely independent behavior without a cause. Under the theory of sociology, i.e., trust increases over time in relationships [28], which means the trust from one node to another node should satisfy the constraints of time. In some social networks such as Facebook and Twitter, an explicit link implies that two nodes are very close for their frequent communication. However, in a trust network, two nodes may be connected but the link may be untrustworthy. More importantly, a trustworthy link in trust network is the situation where two connected nodes have proper time interval. If two users are trustworthy in terms of their time intervals, then their trust links are more trustworthy. For instance, users have trust relations as Fig. 1, where $t_A < t_B < t_C < t_D < t_E$ and $t_1 < t_2 < t_3 < t_4 < t_5 < t_6 < t_7 < t_8$. It is intuitive that the trust link from B to A on t_1 is more trustworthy than that from C to E on t_5. Because E's registration time is later than C's, old node seldom actively trusts new node in general. In contrast, B's registration time is later than A, so B trusts A firstly and then A trusts B as a feedback, which
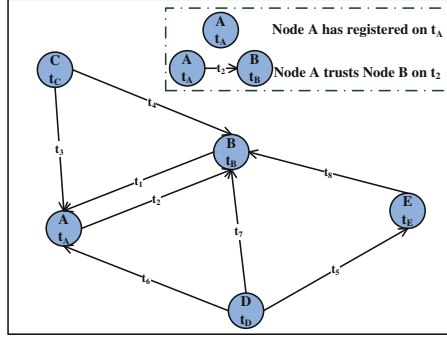
Fig. 1. An example of trust network

is trustworthy. D is a suspicious node used to promote E by connecting other trustworthy nodes such as A and B.

The contribution of this paper include: (1) a novel temporal method (T-PR) which builds upon PageRank and captures how the structure of the interaction network changes over; (2) a new approach for aggregating scores from each of the temporal features over which T-PR is running, the first based on build-up time-length, the second based on frequency, and the third based on similarity; (3) an extensive experiment on a real world data set to evaluate the performance of our proposed method T-PR. Results show that the T-PR outperforms both the state-of-the-art ranking models, and can improve the performance of user recommendation.

The rest of this paper is organized as follows. We briefly describe the PageRank algorithm and review some related work in Sect. 2. In Sect. 3, we define measures of temporal factors, and then describe how to unify them in a Temporal PageRank (T-PR). In Sect. 4, we show and discuss the results of experiments. Finally we make a conclusion and present some future researches in Sect. 5.

2   Related Work

2.1   Brief Review of PageRank

The basic definition of PageRank [25] can be stated as follows. Given a directed graph $G = (V, E)$, while V is the set of nodes and $E \subset V \times V$ is the set of links/edges, without multiple edges. If $u \in V$ has a link to $v \in V$, it implies that u implicitly confers some importance to v. Let $pr(u)$ be the PageRank score of u and $w(u, v)$ be the proportion of importance propagated from u to v, which is normally set to $1/|d^o(u)|$, where $|d^o(u)|$ is the out-degree of u in G. Therefore, link $(u, v) \in E$ confers $pr(u)/|d^o(u)|$ units of rank to v. The PageRank score can be calculated by the following equation:

$$\forall v \in V, pr^{(\phi+1)}(v) = \sum_{u \in d^i(v)} w(u, v)pr^{(\phi)}(u) = \sum_{u \in d^i(v)} \frac{pr^{(\phi)}(u)}{|d^o(u)|} \qquad (1)$$

where $d^i(v)$ represents the set of nodes linking to v. The total amount of score conferred on v is the summation of the score of each u divided by its out-degree.

2.2   Other Related Work

Since the PageRank [25] and HITS [17] algorithms were published, there are a large number of papers on improvements , variations, and speed-up of the algorithms have proposed in [4,6,16,23]. These works are still within the framework of the original algorithms, and do not consider the temporal aspect.

There are also some studies that concentrate on incorporate temporal factor into the PageRank algorithm. [2] proposes a modified PageRank by weighting pages with an exponential decay function according to their age, which is a limited attempt to tackle the problem without evaluation. TimedPageRank (TimedPR) [30] is to weight each citation by an exponential decay function according to publication time of the papers, which cannot be directly applied to social networks. Scientific papers have static information fixed a publication time. Since articles cannot be deleted, their citation counts are monotonically increasing. By contrast, social networks can be modified by adding/deleting links. And [3] presents Time-Aware Authority Ranking (T-Rank), by weighting the page transition and random jump probabilities. Their results may not be representative since each domain usually has its own pattern. [7,10,15,21] are enhanced PageRank with time. However, these approaches are not suitable for social networks; user's behavior on social networks is very different from web pages. In summary, all the researchers of the above studies on (time-)weighted PageRank show their methods are better than the standard PageRank, but neglect a high correlation of various PageRank variants and other social network measures such as build-up time-length, frequency, and similarity. None of the studies, however, has combined time information from both the nature of social networks and the characteristic of users via the "Temporal" PageRank described in this paper.

Recently, several researchers, such as [5,8,20,26,29], have applied machine learning techniques to train the ranking model using queries previously and relevance information on retrieved results derived from user's browser behavior, to improve ranking quality. The success of these "learning to rank" approaches depends on both query and document information, which is very difficult to tune the parameter of theses machine learning techniques.

3   Temporal PageRank

In this section, we first describe quantitative measure of three temporal factors. To evaluate the usefulness of these factors, we then present how to incorporate all three factors into the personalized PageRank algorithm.

3.1   Temporal Aspects

Social network is continually changing, i.e., users are registered, modified, or deleted over time. Based on the hypothesizes: (1) the longer interval between

registration time of user and creation time of link is more trustworthy; (2) trustworthy users are often active and regularly adding/deleting links, and seldom perform majority operations at once; (3) recent link is more important than old link. Let us consider a social network as a directed graph $G = (V, E, T)$, where the nodes represent users with registration time, the edges represent the trust links with creation time. For $\forall u, v \in V$, $t_u, t_v \in T$ is the registration time of u, v respectively, and $\exists (u, v) \in E$, $t_{uv} \in T$ is the creation time of link from u to v.

Built-Up Time-Length Factor. We define built-up time-length factor to describe the interval between registration time of node and creation time of link. A longer time-length is considered to be more trustworthy than a shorter one. Time-length is used to describe the length of the time that something continues or exists.

Let t_{uv} - t_u denote the built-up time-length of u adds a link to v, obviously $t_u \leq t_{uv}$. The Built-up Time-length Factor of (u, v) $BTF(u, v)$, is simply defined in a exponential time-scale as:

$$BTF(u,v) = \begin{cases} 1 - e^{-\frac{(t_{uv} - t_u)}{2\sigma^2}} & if \ t_u \geq t_v, \\ 1 - e^{-\frac{(t_{uv} - t_{vu})}{2\sigma^2}} & if \ t_u < t_v \ and \ \exists t_{uv} \geq t_{vu}, \\ 0 & otherwise. \end{cases} \quad (2)$$

Note that $\sigma = 14$ means fortnight, which is the best empirical setting. This built-up time-length factor is bound by $[0, 1)$. Given the observation period t_{uv} through t_u, the built-up time-length factor of u has value 0 if the registration time of node u is the same as the creation time of link (u, v), increasing to 1 for existing interval between t_{uv} and t_u/t_{vu}. If $t_u < t_v$ and $\exists t_{uv} > t_{vu}$, which mean u is older than v and (u, v) is a feedback of (v, u), then $BTF(u, v) \geq 0$.

Frequency Factor. Let $t_i(u)$ is the i-th timestamp of node u adding links, $N_{t_i(u)}$ is the count of links added by node u on time t_i. n_u is the number of timestamps when u adds links. $\overline{N_{t(u)}} = \frac{\sum_{i=1}^{n} N_{t_i(u)}}{n_u}$ is the average of $N_{t(u)}$, and the variance is $var(u) = \frac{\sum_{i=1}^{n} (N_{t_i(u)} - \overline{N_{t(u)}})^2}{n_u}$. n_u is bigger means u is more active. $var(u)$ is smaller means u is more trustworthy, otherwise, $var(u)$ is bigger means u is less trustworthy. In Fig. 2, node D is more trustworthy than other nodes, since node D is an active user and adds trust links regularly, which satisfies the $var(D)$ is smallest and n is biggest. Frequency factor (FF) can be computed by Eq. 3.

$$FF(u, v) = e^{-\frac{var(u) + N_{t_{uv}(u)}}{n_u \cdot N_{t(u)}}} \quad (3)$$

The frequency factor of (u, v) is the proportion of trustworthiness of u adding a link to v. This frequency factor is bound by $(1, 0)$.
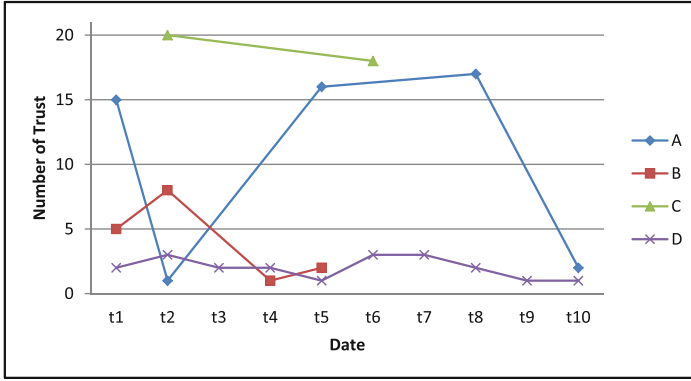
Fig. 2. An example of diverse frequency for adding trust

Similarity Factor. The Built-up Time-length and Frequency factors only consider the temporal factors of individual users. We hypothesize that the pattern of user adds trust link in other causes may also have significance in estimating a user's importance. A user that adds trust links in accordance with a special or good reason (e.g., interest, personality, etc.) reflects trustworthiness in itself, which has been verified experimentally in [12]. However, considering only similarity between all users without examining their ephemeral of behavior may be biased and unfair to those users who have added trust links recently. During a period of time, one user tends to show interest in similar things, such as focusing on similar users. So it is acceptable that an old user trusts a new user because they are similar during a period of time.

A similarity factor describes how similar change the behavior of a user is to others in the network, based on the hypothesis that authoritative users would behave in a similar way [1]. Here, can simply be the SimRank similarity function used in [14]. The similarity factor of u and v on time t is thus considered as the similarity of their in-links and out-links before t_{uv}. Similarity Factor (SF) can be defined as:

$$SF(u,v) = Sim_{t_{uv}}(u,v)e^{-\gamma(t-t_{uv})} \tag{4}$$

A key parameter of the above formula is γ which denotes how fast the value of similarity factor decreases over time. If γ is too small, then the similarity factor will have a long effect. Especially, if $\gamma = 0$ then similarity factor does not decrease with time. On the other hand, if γ is too large, then similarity factor has very short effect, getting quickly to 0. Intuitively, γ should be set according to the timescale of the users to reflect how fast real rankings change. Moreover, it should also be related to the number of trust links added per unit of time (e.g., year or month) of the users. In our experiments, we investigate the impact of various values for γ. In our data set, we find $\gamma = 0.05$ gets the best performance.

And $Sim_{t_{uv}}(u,v)$ is computed by Eq. 5. Let $d^o_{t_{uv}}(u)$ denotes the set of all outgoing neighbors from node u before t_{uv}, and likewise, $d^i_{t_{uv}}(u)$ denotes the set

of all incoming neighbors to node u before t_{uv}. $|d^o_{t_{uv}}(u)|$ and $|d^i_{t_{uv}}(u)|$ denote the number of nodes in $d^o_{t_{uv}}(u)$ and $d^i_{t_{uv}}(u)$ respectively.

$$Sim_{t_{uv}}(u,v) = \frac{\lambda \cdot \theta}{|d^i_{t_{uv}}(u)||d^i_{t_{uv}}(v)|} \sum_{p \in d^i_{t_{uv}}(u)} \sum_{q \in d^i_{t_{uv}}(v)} Sim_{t_{pq}}(p,q)$$
$$+ \frac{\lambda \cdot (1-\theta)}{|d^o_{t_{uv}}(u)||d^o_{t_{uv}}(v)|} \sum_{p \in d^o_{t_{uv}}(u)} \sum_{q \in d^o_{t_{uv}}(v)} Sim_{t_{pq}}(p,q) \tag{5}$$

Here, λ is a decay factor between 0 and 1, $\theta \in [0,1]$ is a parameter, which determining the importance of in-links and out-links in the graph. In this paper, we take $\lambda = 0.8$ refers to [14], and $\theta = 0.5$, which is the best empirical setting.

3.2 Temporal Ranking

In this section, we describe a new ranking function that incorporates the built-up time-length, frequency, and similarity factors into PageRank. Let $w_t(u,v)$ be the weight for a transition from node u to node v on time t, $d^i_t(v)$ denotes the set of all incoming neighbors to node v before t. Then, the PageRank formula in Eq. 1 can be rewritten in a temporal form (T-PR) as follows:

$$\forall v \in V, tpr^{(\phi+1)}(v) = \sum_{u \in d^i_t(v)} w_t(u,v)tpr^{(\phi)}(u) \tag{6}$$

In the original PageRank, $w_t(u,v)$ is set to $1/|d^o(u)|$. Some studies [18,27] have proposed biased PageRank. In this paper, our purpose is primarily to investigate the contribution of temporal information (i.e., built-up time-length, frequency, and similarity) to node ranking by weighting in node transitions, and subsequently to reduce the effect of bias against new-born nodes by weighting nodes with only link information.

The time-dependent interactions among users clearly have a fundamental impact on their rankings. For example, a series of recent trust links is likely to be more important in determining the current ranking of a set of users than a series of trust links among the same set of users that occurred far in the past. Intuitively, the importance of trusts link for ranking users decays over time: the older a trust link is, the less important is its result. Thus, a good ranking for todays users should give more importance to recent trust links.

The importance of relationships is usually captured by assigning weights to network edges. Thus, based on this intuition, we propose the use of temporal edge weights to reflect the fact that the importance of trust links decays with time. In particular, we will consider an exponentially decaying weight, controlled by a parameter that determines how fast importance decays over time. Consider two nodes u and v. Moreover, let $w_t(u,v)$ denote the weight of the directed edge from u to v on time $t \geq 0$. We define the edge weight $w_t(u,v)$ as follows:

$$w_t(u,v) = \frac{\alpha BTF(u,v) + \beta FF(u,v) + (1-\alpha-\beta)SF(u,v)}{\sum_{s \in d^o_t(u)} \alpha BTF(u,s) + \beta FF(u,s) + (1-\alpha-\beta)SF(u,s)} \tag{7}$$

Note that $\alpha \in [0, 1]$ is a parameter used to weigh the importance of BTF and $\beta \in [0, 1]$ is a parameter used to weigh the importance of FF. In this paper, we take $\alpha = \beta = 0.3$, which is the best empirical setting.

4   Evaluation Methodology

We implemented five ranking algorithms: PR [25], TWPR [21], T-Rank [3], TimedPR [30], and T-PR (new) for comparison of the quality of ranking results.

4.1   Experimental Data Set

As the source of web data, we selected a portion of the set of trust network from Epinions.com. To study the evolution of trust network, we downloaded them until April 6th, 2010. The data include the user ID with registration date and the trust link from one user to another user with creation date. Because the creation date of trust link is unavailable before January 1st, 2001, our experimental data only include the users having trust link after January 1st, 2001. Table 1 shows the various classes of statistics about Epinions data. Figure 3 describes the evolution of registered nodes and trust links per half a year. Figure 4 displays the count of users with different in-degrees on April 6th, 2010.

Table 1. Statistical information of Epinions trust network

Epinions	Number	From	To	Total days	Avg
Nodes	47109	25-Jun-1999	23-Mar-2010	3768	12.502
Trusts	258966	01-Jan-2001	02-Apr-2010	3359	77.096
Trusts/Nodes:	5.497	Nodes trust:	33875	Nodes trusted:	29807

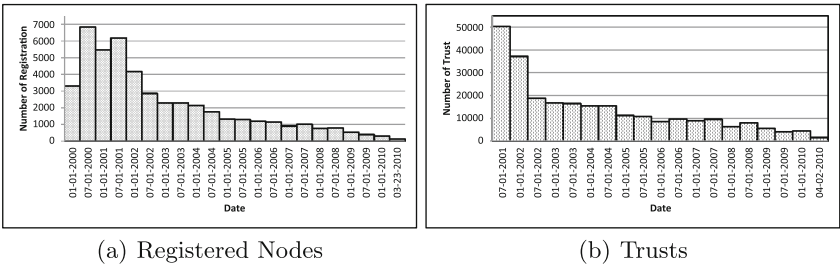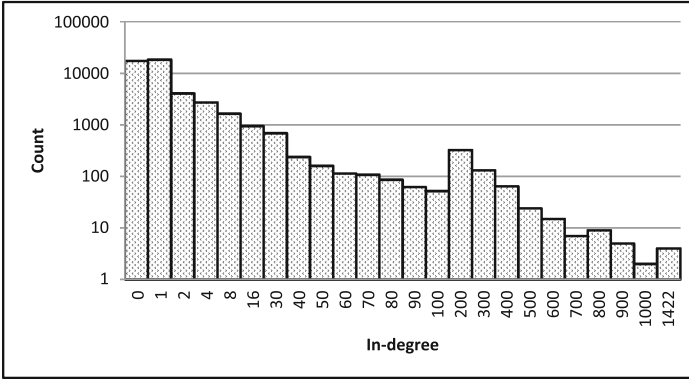(a) Registered Nodes   (b) Trusts

Fig. 3. The evolution of Epinions trust network

Fig. 4. Distribution of in-degree for Epinions

4.2 Evaluation Methods

In this section, we detail the three measures used in this paper. The $OSim$ and $KSim$ metrics, proposed by Haveliwala [11], measure the similarity of any two ranked lists l_1 and l_2, each of size k. $OSim(l_1, l_2)$ determines the degree of overlap between the top-k nodes of two rankings l_1 and l_2.

$$OSim(l_1, l_2) = \frac{|l_1 \cap l_2|}{k} \qquad (8)$$

where l_1 and l_2 are the lists of top-k nodes.

$KSim(l_1, l_2)$ determines the degree of agreement in which the relative ordering of the top-k nodes of two lists l_1 and l_2, have the same relative order in both rankings. Consider two lists l_1 and l_2 of top-k rankings. Let U be the union of nodes contained in both lists and define l_1' as the extension of l_1 to add the elements $U - l_1$ at l_1''s end. Similarly, l_2' is also defined as the extension of l_2. We can define $KSim$ as follows:

$$KSim(l_1, l_2) = \frac{|(u, v) : l_1', l_2' \ agree \ on \ order \ of \ (u, v) \ and \ u \neq v|}{|U| \times (|U| - 1)} \qquad (9)$$

where the numerator denotes the number of pairwise agreements of elements between l_1' and l_2'.

The final measure we used to evaluate the performance of these 5 algorithms is the Mean Reciprocal Rank (MRR). Given an ordered list of predicted nodes (these nodes are in a descending order according to their ranking values) and the "actual node" that is trustworthy, the reciprocal rank is calculated as $1/\rho$ where ρ is the position of the "actual node" in this ordered list, and it is set to 0 if the "actual node" is not in this list. The MRR value of an algorithm that is closer to 1 denotes a better performance.

4.3   Results and Discussions

We conducted experiments on our crawled data to study the node authority assessment of the PR, TWPR, T-Rank, TimedPR, and T-PR (new) methods. In the following, we first discuss the time evolution of authoritativeness. Then, we describe results of a user study to evaluate the quality of ranked results.

The first set of experiments measure the comparison of different scores of the nodes, which compares the ranking of nodes using the different scores against that produced by the PageRank [25] on January 1st, 2010. Figure 5 shows the comparison of results. Note that we have scaled the scores by multiplying with 1000. The distribution of T-PR is steepest due to the consideration of the effect of 3 temporal factors. One common trend we observe is that nodes with less PR have low score, and those with high PR have high score. This shows that the ranking determined by the T-PR conform to the perception that more popular nodes have more trustworthy.

For each timestamp, we first computed an authoritative score for each node according to a ranking method, and then sorted them from highest to lowest to produce a ranked list. To investigate the time evolution effect (i.e., how the authoritativeness changes over time), we compared the lists of top authoritative nodes obtained from those 5 ranking methods for several consecutive half a year. Figure 6 shows the evolution of top-500 authorities in terms of OSim and KSim. Note that the observation period starts from January 1st, 2002 until April 2nd, 2010. Hence, each point on the graph is the similarity between the ranking at that time and that of the previous half a year.

As shown in the Fig. 6, T-PR yields OSim values of 0.819-0.845 (average 0.831) and KSim values of 0.822-0.851 (average 0.837). This implies that PR produces nearly the same rankings over the entire observation period on the experimental data set. In contrast, the consecutive rankings of TWPR, T-Rank, TimedPR, and T-PR have lower similarities because of their temporal approach. We also see that T-PR produces more similar rankings than TWPR, T-Rank
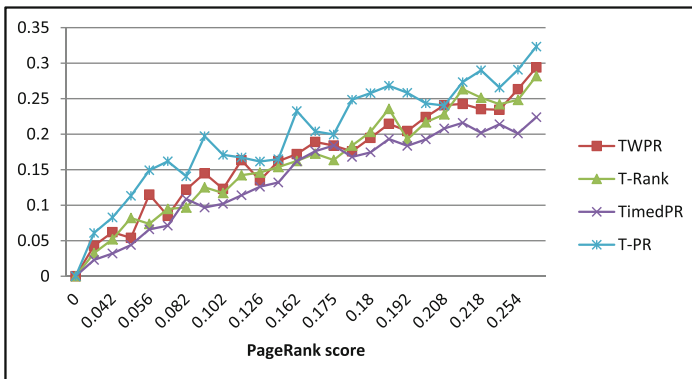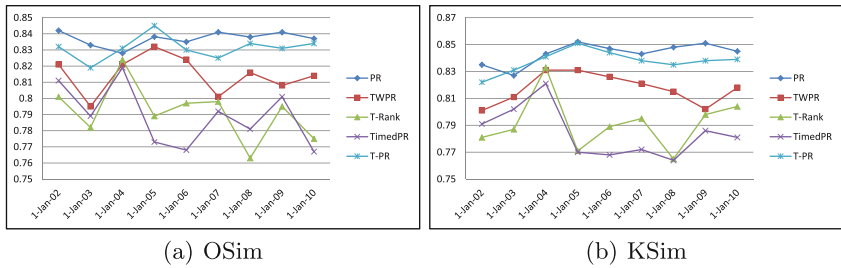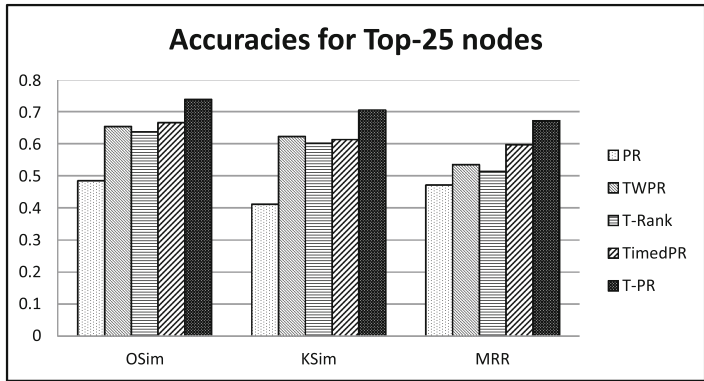

Fig. 5. Comparison of different scores with PR for Epinions

(a) OSim   (b) KSim

Fig. 6. Similarities between lists of top-500 authoritative nodes produced by the same ranking methods for consecutive years
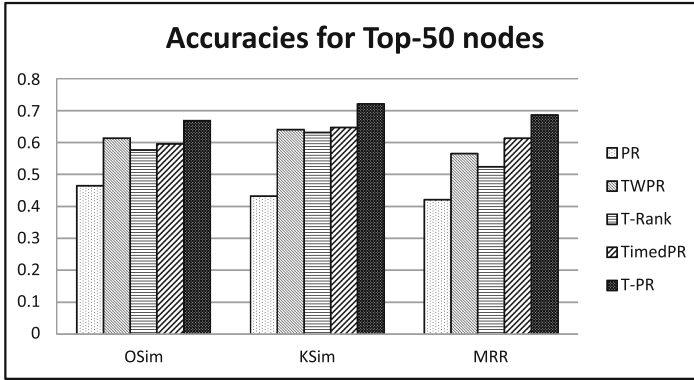

Fig. 7. Average OSim, KSim and MRR values for top-25 nodes

and TimedPR, respectively. This phenomenon may come from the effect of the BTF and FF factors in T-PR that allow some authoritative nodes in the past to be ranked in the top authorities.

We used these 5 algorithms (PR, TWPR, T-Rank, TimedPR, and T-PR) to obtain the sets of the top-25 and top-50 most highly-ranked nodes for the most popular nodes we chose from the data set. Then for each of the most popular nodes, we compared these sets of the top-25 and top-50 nodes with the sets of the top-25 and top-50 most popular nodes from Epinions.com respectively by using the aforementioned evaluation measures. Figures 7 and 8 show the average OSim, KSim and MRR values for the top-25 and top-50 rankings of these 5 different algorithms.

As depicted in Figs. 7 and 8, T-PR outperforms PR, TWPR, T-Rank and TimedPR in all OSim, KSim and MRR values. In the top-25 ranking, T-PR outperforms other methods by 10.8%, 13.1% and 12.5% for OSim, KSim and MRR respectively. In the top-50 ranking, the corresponding percentages are 8.9%, 11.4%, 11.7% respectively. Therefore, we can confirm that adopting the built-up time-length, frequency, and similarity of a node adding trust link can improve the accuracy of nodes ranking.

Fig. 8. Average OSim, KSim and MRR values for top-50 nodes

Table 2. Top-20 users' different scores in Epinions

Name	Registration date	PR	TWPR	T-Rank	TimedPR	T-PR
jo.com	02/23/2000	0.285	0.325	0.313	0.297	0.341
dkozin	10/04/1999	0.256	0.293	0.267	0.281	0.323
mkaresh	10/16/1999	0.262	0.301	0.255	0.266	0.305
Freak369	02/05/2000	0.277	0.256	0.261	0.243	0.298
Bryan_Carey	12/21/1999	0.214	0.272	0.234	0.258	0.292
three_ster	06/26/2000	0.226	0.207	0.210	0.213	0.286
shoplmart	11/26/2001	0.265	0.197	0.187	0.221	0.281
dlstewart	05/31/2002	0.244	0.211	0.196	0.206	0.264
Howard_Creech	08/16/1999	0.229	0.174	0.182	0.195	0.249
ChrisJarmick	07/19/2000	0.212	0.203	0.173	0.188	0.243
popsrocks	08/25/2002	0.221	0.199	0.164	0.175	0.233
mrkstvns	09/09/1999	0.208	0.215	0.158	0.191	0.210
yusakugo	04/14/2000	0.258	0.178	0.152	0.162	0.202
melissasrn	02/03/2001	0.211	0.164	0.171	0.153	0.194
surferdude7	01/08/2000	0.188	0.193	0.163	0.147	0.186
AliventiAsylum	08/11/2000	0.174	0.161	0.144	0.168	0.181
jeremy1456	03/17/2001	0.166	0.186	0.155	0.156	0.168
marytara	04/02/2001	0.171	0.157	0.149	0.144	0.157
captaind	01/19/2003	0.175	0.151	0.133	0.132	0.152
yakkowarner	03/01/2004	0.149	0.146	0.143	0.135	0.141

4.4   Case Study

In the final experiment, we want to compare the scores computed by 5 differ-ent algorithms including PR, TWPR, T-Rank, TimedPR, and T-PR(new) for

Top-20 trustworthy users from Epinions.com. We select Top-20 reviewers for the most popular authors overall obtained from Epinions.com. In Table 2, we can find that our method produces a better result than others. The T-PR scores conform to the manual ranking. All scores in Table 2 is multiplied by 1000. From the Table 2, we find the PR, TWPR, T-Rank, and TimedPR values fluctuate erratically, and our method get better result than others, which proves Eq. 6 is an effective improvement of PageRank.

5   Conclusion

In this paper, we propose a novel Temporal PageRank algorithm on social networks. Three temporal factors (built-up time-length, frequency, and similarity) are used to improve PageRank so that it favor the nodes that were registered for a longer time, more frequently, and recent trust link weighed more than others. We conduct extensive experiments to evaluate the performance of T-PR. Evaluation results show that the proposed T-PR outperforms other timed PageRank algorithms significantly. Moreover, our T-PR algorithm provides most similar rankings to human beings' preference. In our future work, we are interested in studying the effect of the time factor using other decay functions. Since social networks are multi-relational networks, we will incorporate more relations to improve the ranking of Temporal PageRank.

References

1. Anagnostopoulos, A., Kumar, R., Mahdian, M.: Influence and correlation in social networks. In: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 7–15. ACM (2008)
2. Baeza-Yates, R., Saint-Jean, F., Castillo, C.D.: Web structure, dynamics and page quality. In: Laender, A.H.F., Oliveira, A.L. (eds.) SPIRE 2002. LNCS, vol. 2476, pp. 117–130. Springer, Heidelberg (2002)
3. Berberich, K., Vazirgiannis, M., Weikum, G.: Time-aware authority ranking. Internet Math. **2**(3), 301–332 (2005)
4. Borodin, A., Roberts, G.O., Rosenthal, J.S., Tsaparas, P.: Finding authorities and hubs from link structures on the world wide web. In: Proceedings of the 10th International Conference on World Wide Web, pp. 415–429. ACM (2001)
5. Burges, C., Shaked, T., Renshaw, E., Lazier, A., Deeds, M., Hamilton, N., Hullender, G.: Learning to rank using gradient descent. In: Proceedings of the 22nd International Conference on Machine learning, pp. 89–96. ACM (2005)
6. Cho, J., Roy, S.: Impact of search engines on page popularity. In: Proceedings of the 13th International Conference on World Wide Web, pp. 20–29. ACM (2004)

7. Fiala, D.: Time-aware PageRank for bibliographic networks. J. Informetrics **6**(3), 370–388 (2012)
8. Geng, X., Liu, T.Y., Qin, T., Arnold, A., Li, H., Shum, H.Y.: Query dependent ranking using k-nearest neighbor. In: Proceedings of the 31st aNnual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 115–122. ACM (2008)
9. Gonçalves, B., Meiss, M.R., Ramasco, J.J., Flammini, A., Menczer, F.: Remembering what we like: toward an agent-based model of web traffic (2009). arXiv preprint arXiv:0901.3839
10. Guo, Y.Z., Ramamohanarao, K., Park, L.A.: Personalized pagerank for web page prediction based on access time-length and frequency. In: IEEE/WIC/ACM International Conference on Web Intelligence, pp. 687–690. IEEE (2007)
11. Haveliwala, T.H.: Topic-sensitive PageRank: a context-sensitive ranking algorithm for web search. IEEE Trans. Knowl. Data Eng. **15**(4), 784–796 (2003)
12. Hu, W., Gong, Z.: Assessing the credibility of nodes on multiple-relational social networks. In: Benatallah, B., Bestavros, A., Manolopoulos, Y., Vakali, A., Zhang, Y. (eds.) WISE 2014, Part II. LNCS, vol. 8787, pp. 62–77. Springer, Heidelberg (2014)
13. Ieong, S., Mishra, N., Sadikov, E., Zhang, L.: Domain bias in web search. In: Proceedings of the Fifth ACM International Conference on Web Search and Data Mining, pp. 413–422. ACM (2012)
14. Jeh, G., Widom, J.: Simrank: a measure of structural-context similarity. In: KDD, pp. 538–543 (2002)
15. Júnior, P.S.P., Gonçalves, M.A., Laender, A.H., Salles, T., Figueiredo, D.: Time-aware ranking in sport social networks. J. Inf. Data Manag. **3**(3), 195 (2012)
16. Kamvar, S.D., Haveliwala, T.H., Manning, C.D., Golub, G.H.: Extrapolation methods for accelerating pagerank computations. In: Proceedings of the 12th International Conference on World Wide Web, pp. 261–270. ACM (2003)
17. Kleinberg, J.M.: Authoritative sources in a hyperlinked environment. J. ACM **46**(5), 604–632 (1999)
18. Kohlschütter, C., Chirita, P.A., Nejdl, W.: Utility analysis for topically biased pagerank. In: Proceedings of the 16th International Conference on World Wide Web, WWW 2007, pp. 1211–1212. ACM, New York (2007). http://doi.acm.org/10.1145/1242572.1242770
19. Liu, Y., Gao, B., Liu, T.Y., Zhang, Y., Ma, Z., He, S., Li, H.: Browserank: letting web users vote for page importance. In: Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 451–458. ACM (2008)
20. Liu, Y., Liu, T.Y., Gao, B., Ma, Z., Li, H.: A framework to compute page importance based on user behaviors. Inf. Retrieval **13**(1), 22–45 (2010)
21. Manaskasemsak, B., Rungsawang, A., Yamana, H.: Time-weighted web authoritative ranking. Inf. Retrieval **14**(2), 133–157 (2011)
22. Meiss, M.R., Menczer, F., Fortunato, S., Flammini, A., Vespignani, A.: Ranking web sites with real user traffic. In: Proceedings of the 2008 International Conference on Web Search and Data Mining, pp. 65–76. ACM (2008)
23. Ntoulas, A., Cho, J., Olston, C.: What's new on the web? The evolution of the web from a search engine perspective. In: Proceedings of the 13th International Conference on World Wide Web, pp. 1–12. ACM (2004)
24. Otterbacher, J., Hemphill, L., Dekker, E.: Helpful to you is useful to me: the use and interpretation of social voting. Proc. Am. Soc. Inf. Sci. Technol. **48**(1), 1–10 (2011)

25. Page, L., Brin, S., Motwani, R., Winograd, T.: The pagerank citation ranking: bringing order to the web. Technical report, Stanford InfoLab, November 1999. http://ilpubs.stanford.edu:8090/422/
26. Richardson, M., Prakash, A., Brill, E.: Beyond pagerank: machine learning for static ranking. In: Proceedings of the 15th International Conference on World Wide Web, pp. 707–715. ACM (2006)
27. Rungsawang, A., Puntumapon, K., Manaskasemsak, B.: Un-biasing the link farm effect in pagerank computation. In: 2007 21st International Conference on Advanced Information Networking and Applications, AINA 2007, pp. 924–931. IEEE (2007)
28. Vanneste, B., Puranam, P., Kretschmer, T.: Trust over time in exchange relationships: meta-analysis and theory. Strat. Manag.J. (2013). Forthcoming
29. Xu, J., Li, H.: Adarank: a boosting algorithm for information retrieval. In: Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 391–398. ACM (2007)
30. Yu, P.S., Li, X., Liu, B.: Adding the temporal dimension to search-a case study in publication search. In: 2005 Proceedings of the 2005 IEEE/WIC/ACM International Conference on Web Intelligence, pp. 543–549. IEEE (2005)