Efficient Calculation of Personalized Document Rankings

Claudia Hess, Klaus Stein
Laboratory for Semantic Information Technology
Bamberg University
{claudia.hess, klaus.stein}@wiai.uni-bamberg.de

Abstract

Social networks allow users getting personalized recommendations for interesting resources like websites or scientific papers by using reviews of users they trust. Search engines rank documents by using the reference structure to compute a visibility for each document with reference structure-based functions like PageRank. Personalized document visibilities can be computed by integrating both approaches. We present a framework for incorporating the information from both networks, and ranking algorithms using this information for personalized recommendations. Because the computation of document visibilities is costly and therefore cannot be done in runtime, i.e., when a user searches a document repository, we pay special attention to develop algorithms providing an efficient calculation of personalized visibilities at query time based on precalculated global visibilities. The presented ranking algorithms are evaluated by a simulation study.

1 Introduction

Recently several approaches for combining social networks and document reference networks have been proposed [Hess *et al.*, 2006], [Stein and Hess, 2006], [García-Barriocanal and Sicilia, 2005], [Korfiatis and Naeve, 2005]. They integrate information on social relationships into classical reference-based measures such as PageRank [Page *et al.*, 1998] or HITS [Kleinberg, 1999]. Trust networks gained much attention because trust relationships constitute a strong basis for personalized recommendations as shown by trust-based recommender systems such as [Golbeck, 2006], [Massa and Avesani, 2004], [Avesani *et al.*, 2005]. We call such measures operating on two-layer networks trust-enhanced visibility measures.[1] They are motivated by the fact that a document can be highly visible although its content is completely untrustworthy. Examples are cases of scientific misconduct in which publications considered as valid or even as landmark papers such as the Science papers by the South-Corean stem cell researcher Hwang[2] are declared as faked. Reference-based measures still assign them a high visibility because citations are rarely removed and, as e.g. [Budd *et al.*, 1998] showed, faked papers continue to be cited even after official retraction. In this case, more accurate rankings can be provided by looking at other users' recommendations: as it is a fake, it will no longer be recommended by anyone knowing this. Less extreme, though more frequent is that opinions on the same document differ greatly between users, for example, when a user has a very extreme, or a very progressive opinion. So this user might consider a document as very interesting, whereas most users deem it as untrustworthy. Here, it is crucial whose recommendations are taken into account: considering document reviews depending on the user's trust in the reviewer highly personalizes recommendations. However, as it is reasonable to assume that only a small fraction of documents is reviewed, we integrate trust-weighted reviews into reference-based measures that calculate recommendations for all documents. We therefore have a two-layer architecture with a document reference network and a reader trust network being connected by reviews.

Current recommender systems do not yet integrate this information although parts are already available on the web. A user's bookmarks made available via applications such as del.iciou.us[3] are e.g. simple reviews of webpages. The number of trust networks on the web increases, too. Well-known applications are Epinions or communities such as Orkut. Many users provide FOAF (Friend-of-a-Friend) files (see e. g. [Dumbill, 2002]) with their profile and relationships. An extension of the FOAF vocabulary encodes trust information[4].

This paper analyzes how information from trust and document networks can be integrated into algorithms for personalized recommendations. Based on a general framework for

[1] Other approaches like TrustRank [Gyöngyi *et al.*, 2004] directly attach reliability information to a subset of documents to improve recommendations, which does not correspond to the notion of social trust (derived from a trust network) and does not allow for personalization.

[2] E.g. news@nature.com: http://www.nature.com/news/2005/051219/full/051219-3.html, (accessed June 28, 2006).

[3] http://del.icio.us/

[4] See the ontology for trust ratings at http://trust.mindswap.org/ont/trust.owl (accessed June 29, 2006)

such trust-enhanced visibility measures, we develop concrete functions. To consider up-to-date trust information and reviews, these functions have to be efficiently computable at query time, i. e., in the moment a user searches a document repository. As it is the reference-based measure that is typically very costly, we explore how to use precalculated visibilities. The efficiency of the measures introduced is analyzed with respect to recommendations for a single document and rankings of a set of documents. In the scope of a simulation study, the results obtained by the different functions are compared. The rest of the paper is hence structured as follows: Section 2 discusses the general framework. In section 3, we develop different trust-review-enhanced visibility measures. Section 4 presents the simulation study in which the functions are evaluated and section 5 gives the conclusion.

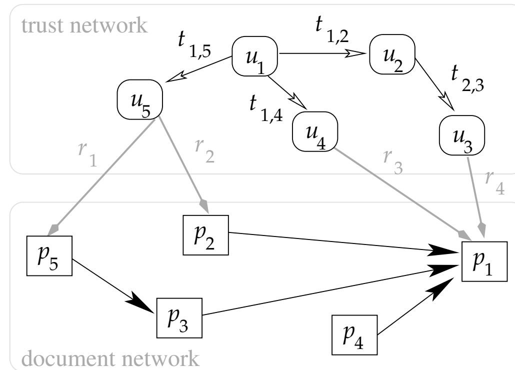

Figure 1: Trust and Document Reference Networks

2 Framework for a Trust-Enhanced Visibility

2.1 Trust and Document Reference Networks

The two-layer architecture encompasses a document reference network and a trust network as shown in figure 1. Documents such as webpages or scientific papers refer to other documents via hyperlinks or citations. Based on the reference structure, a visibility can be calculated for each document, i.e., its importance or rank. The best-known ranking algorithm is PageRank that has originally been incorporated in Google. It computes the visibility vis_d of a document p_d by using the weighted sum of the visibilities vis_k of the papers p_k citing p_d (based on the idea that a paper cited from many important papers must be somehow important)[5]:

$$\text{vis}_{p_d} = \frac{1-\alpha}{N} + \alpha \sum_{p_k \in B_d} \frac{\text{vis}_{p_k}}{|C_k|}$$

where B_d is the set of pages citing p_d and C_k is the set of pages cited by p_k.[6] Originally N is the number of documents in the network, in general it is simply a linear scaling factor. An important feature of this function is that the visibility of any document p_d depends on the visibilities of documents $p_j \in B_d$ citing it. Therefore, changing the visibility of one document influences the visibility of other documents. Other approaches, e. g. HITS, which determines a hub and an authority value for webpages, are based on such a recursive definition, too. The framework presented works with any reference structure-based visibility measure.

The trust network is established between reviewers, i.e., readers or editors who express their opinion on documents. Reviewers assign a trust value to other reviewers, giving directed and weighted edges in the trust network. High trust means that the evaluating user appreciates the evaluated user's reviews, e. g., because he or she applies similar criteria in the review. Trust between indirectly connected users, i.e., the "friends of our friends", can be inferred by trust propagation: if user u_m has not directly evaluated u_n, the trust value $t_{u_m \to u_n}$ for u_n is derived by aggregating the trust values on the path from u_m to u_n. Examples for trust metrics include [Golbeck *et al.*, 2003] and [Avesani *et al.*, 2005].[7] Reviews connect trust and document networks.

2.2 Personalizing Recommendations

The two-layer architecture permits calculating a recommendation for a document (page) p_d from the perspective of a person (user) u_m. For the recommendation, a citation-based measure on the document reference network is enhanced with information from the trust network. As this measure considers in addition the reviews made by the users in the trust network, it is called a trust-review-enhanced visibility, in the following: tre-visibility ($\text{vis}^{\text{tre}}_{p_d}$).

By interpolation on the trust network, the reviews of other users can be taken into account. So we are able to calculate recommendations for documents that are not reviewed by the requesting user her/himself. The information from the trust network tells us to which degree a review should influence the recommendation $\text{vis}^{\text{tre}}_{p_d}$: the reviews made by users who user u_m considers as trustworthy should have most impact. So all reviews are personalized by the user's view on the trust network. By interpolating on the document reference network, recommendations cannot only be calculated for documents that have been reviewed but for all documents. As the visibility of a document depends on the visibilities of the documents citing it, reviews have an indirect impact on adjacent documents: in the example from Fig. 1, reviews r_3 and r_4 will influence the visibility of p_1 if u_1 considers the reviewers u_3

[5]The basic idea of this algorithm was used before by [Pinski and Narin, 1976] to compute the importance of scientific journals.

[6]For n pages this gives a linear system of n equations. Solving this equation system is possible but (for large n) very expensive, so an iterative approach is used. First all vis_i are set to some default value and then the new values r'_i are calculated repeatingly until all vis_i converge (for a discussion of convergence problems in leaves see [Page *et al.*, 1998]).

[7]As we aim to personalize recommendations, we use one of these trust metrics providing trust values from the user's perspective. Other trust metrics calculate global values, e. g., with metrics in the style of PageRank.

and u_4 as trustworthy. While the trust $t_{u_1 \to u_4}$ of u_1 in u_4 is directly given, $t_{u_1 \to u_3}$ is interpolated by some trust metric (see Sec. 2.1) by $t_{u_1 \to u_2} \circ t_{u_2 \to u_3}$. The reviews r_2 and r_1 exert an indirect influence on $\mathrm{vis}^{\mathrm{tre}}_{p_1}$ through the ciations from p_2 and p_5 to p_1.

2.3 The Trust-Enhanced Visibility

We now design the tre-visibility-framework using the structure of the document network and the trust-weighted reviews as the basis for personalizing document visibilities. In the first step by interpolation on the trust network, the trust of a user u_m to all others reachable from her or him is computed. Non-reachable users are given a default trust t_{default} (setting $t_{\mathrm{default}} = 0$ implies that the corresponding reviews have no impact). This interpolation can be done efficiently for all users in parallel. For all reviews r_j of a user u_n, we consider u_m's trust $t_{u_m \to u_n}$ in u_n to be the trust in all reviews by u_n. Now u_m's trust in every review is known. In the second step, personalized tre-visibilities of all documents are computed by incorporating the reviews and weighting them by the trust u_m has in them. Here we can choose between two different approaches:

1. compute a document base visibility $\mathrm{vis}^{\circ}_{p_j}$ of all documents p_j using some visibility function like PageRank and then derive the personalized tre-visibility $\mathrm{vis}^{\mathrm{tre}}_{p_j}$ from the user-independent document base visibility $\mathrm{vis}^{\circ}_{p_j}$ by including trustworthy reviews, or
2. use a modified visibility function that incorporates the reviews on each document directly when computing $\mathrm{vis}^{\mathrm{tre}}_{p_j}$.

The first approach has the advantage to be simple and to be able to precompute[8] $\mathrm{vis}^{\circ}_{p_j}$ for all documents as $\mathrm{vis}^{\circ}_{p_j}$ is user-independent, but the integration of indirect reviews is not straightforward. The second approach automatically handles indirect reviews because the tre-visibilities are used for propagation, but here everything has to be computed on the fly, because no user-independent part exists. To be able to precompute most of the values is important for providing personalized recommendations in search engines to queries such as "should I buy this pay-per-view?" and for sorting query results in a personalized ranking.[9] Anything that has to be computed at query time increases the load of the document repository server and demands the user to wait.

Regardless of the approach used, a tre-visibility function should satisfy the following properties:

1. A review's impact on a document recommendation depends on the degree of trust that the requesting user has in the reviewer. Reviews provided by users who are fully trusted should have a considerable impact on the recommendation, whereas reviews by users deemed as untrustworthy should have minimal impact.
2. If no review by a trustworthy person is available, the recommendation will consist of the visibility calculated on the document network. Although this mere visibility measure is not personalized, it is appropriate: having no review by a trustworthy user does not permit inferring that the document is not worth reading.
3. The degree of influence that reviews have compared with the pure structure-based visibility should be adjustable.
4. Trust-weighted reviews of a document p_j exert an indirect influence on the visibilities of the papers referenced in p_j, because p_j's visibility is modified by the reviews, and so it propagates a modified visibility to the documents that it cites.

The reference-based visibility measure has to be chosen depending on the type of document network: cyclic networks such as the Web normally require different measures as typically acyclic publication networks.

3 TRE-Visibilities

In this section we introduce several functions to compute the trust-review-enhanced document visibility. We use the following definitions: for a document p_d, R_{p_d} is the set of direct reviews on p_d, and $\mathrm{vis}^{\circ}_{p_d}$ is the (precalculated) document base visibility. The distance $k_{i,d}$ of a direct review $r_i \in R_{p_j}$ to p_d is the length[10] of the shortest path from p_j to p_d (if no path exists, $k_{i,d} := \infty$).

All trust values are in $[0,1]$. Reviews are non-negative and in the same range as the visibilities computed by the chosen structure-based visibility function.[11]

3.1 Simple TRE-Visibility

The simplest approach to compute the tre-visibility of p_d is to combine the reviews $r_i \in R_{p_d}$ of p_d, weighted with the respective trust t_i in r_i, and $\mathrm{vis}^{\circ}_{p_d}$. To indicate the impact of $\mathrm{vis}^{\circ}_{p_d}$, it is weighted by its visibility contribution vc which is globally set by the user, e. g., $\mathrm{vc} := 0.5$. So the document base visibility can be treated as additional review with $r_0 := \mathrm{vis}^{\circ}_{p_d}$ and $t_0 := \mathrm{vc}$. This gives the simple tre-visibility:

$$\mathrm{vis}^{\mathrm{tre}_s}_{p_d} = \frac{\sum_{i=0}^{n} t_i r_i}{\sum_{i=0}^{n} t_i} = \frac{\mathrm{vc} \cdot \mathrm{vis}^{\circ}_{p_d} + \sum_{i=1}^{n} t_i r_i}{\mathrm{vc} + \sum_{i=1}^{n} t_i} .$$

---

[8] as Google does to rank millions of webpages

[9] As users normally only read the documents provided on the first page of the result listing this ranking is fairly important.

[10] i. e. the number of edges

[11] This can be achieved by scaling either the reviews or the visibilities, e. g., by choosing an appropriate N in the PageRank.

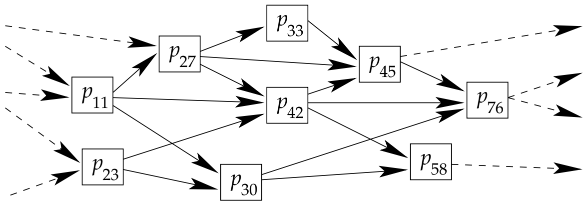

Figure 2: part of a document reference network

3.2 Integrated TRE-Visibility

The tre$_s$-visibility algorithm is appealing in its simplicity but neglects indirect reviews. As in structure-based visibility algorithms like PageRank or HITS the visibility of a document depends on the visibilities of the documents citing it, an indirect effect of reviews on adjacent documents can be achieved by simply swapping the computation sequence: instead of first computing the document visibility and then adding the reviews, the reviews are directly incorporated in the visibility function, e. g. with PageRank[12]:

$$
\begin{aligned}
\mathrm{vis}'_{p_d} &= \frac{1-\alpha}{N} + \alpha \sum_{p_k \in B_d} \frac{\mathrm{vis}^{\mathrm{tre}_i}_{p_k}}{|C_k|} \\
\mathrm{vis}^{\mathrm{tre}_i}_{p_d} &= \frac{\mathrm{vc} \cdot \mathrm{vis}'_{p_d} + \sum_{i=1}^{n} t_i r_i}{\mathrm{vc} + \sum_{i=1}^{n} t_i} .
\end{aligned}
$$

As the visibility $\mathrm{vis}^{\mathrm{tre}_i}_{p_d}$ of a document p_d now depends on the visibilities $\mathrm{vis}^{\mathrm{tre}_i}_{p_k}$ of the documents $p_k \in B_d$ citing it, the reviews on the documents p_k have an impact on $\mathrm{vis}^{\mathrm{tre}_i}_{p_d}$; in other words: the trust-review-enhanced visibilities are propagated through the document network.

The drawback of this approach is that the visibilities of *all* documents are now user-dependent: the visibility vis'_{p_d} of a document p_d can no longer be precomputed offline because it is influenced by the trust-weighted reviews of the documents citing it, and on those of the documents citing these and so on.[13] Precomputing personalized rankings for all users would only be possible for a very small set of users (a small trust network), while computing the complete ranking on the fly at query time is only possible for small document sets. So we need some more efficient visibility algorithm.

3.3 Path-based TRE-Visibility

We therefore have to develop a function that uses precalculated visibilities $\mathrm{vis}^{\circ}_{p_d}$ for all documents p_d like the tre$_s$-visibility function (Sec. 3.1) but that additionally takes into account indirect reviews like the integrated tre$_i$-visibility (Sec. 3.2).

The tre$_i$-visibility $\mathrm{vis}^{tre_i}_{p_{58}}$ of document p_{58} shown in Fig. 2 is computed by the reviews on p_{58} and the visibilities $\mathrm{vis}^{tre_i}_{p_{42}}$ and $\mathrm{vis}^{tre_i}_{p_{30}}$ of p_{42} and p_{30}, which depend on the reviews on p_{42} and p_{30} and so on. The visibility $\mathrm{vis}^{tre_i}_{p_{42}}$ of p_{42} contributes by $\frac{1}{3}$ and $\mathrm{vis}^{tre_i}_{p_{30}}$ by $\frac{1}{2}$ to $\mathrm{vis}^{\circ}_{p_{58}}$, because p_{42} has 3 and p_{30} 2 outgoing citations and (according to PageRank) the visibility of a document is distributed over all outgoing citations. The reviews on p_{42} thus contribute by $\frac{1}{3}$ to p_{58}, as they are part of $\mathrm{vis}^{tre_i}_{p_{42}}$. This also works for larger distances: a review r_i on p_{11} contributes to p_{42} by $\frac{1}{3}$ and therefore along the path $[\ p_{11} \xrightarrow{\frac{1}{3}} p_{42} \xrightarrow{\frac{1}{3}} p_{58}\ ]$ by $\frac{1}{9}$, and along the path $[\ p_{11} \xrightarrow{\frac{1}{3}} p_{30} \xrightarrow{\frac{1}{2}} p_{58}\ ]$ by $\frac{1}{6}$. The total contribution of a review r_i on p_{11} to the visibility of p_{42} is therefore $\frac{1}{9} + \frac{1}{6} = \frac{5}{18}$.

In general: the contribution c^{π}_{j} of a review r_i of a document p_j along a path $\pi = [\ p_j \to q_1 \to q_2 \to \ldots \to q_m \to p_d\ ]$ is $c^{\pi}_{j} = \frac{1}{|C_{p_j}|} + \frac{1}{|C_{q_1}|} + \frac{1}{|C_{q_2}|} + \ldots + \frac{1}{|C_{q_m}|}$.

By restricting the review influence to some maximum distance $k_{\max}$, all indirect reviews and their contributions can be precomputed offline for all documents.[14] So we get

$$
\mathrm{vis}^{\mathrm{tre}_p}_{p_d} = \frac{\mathrm{vc} \cdot \mathrm{vis}^{\circ}_{p_d} + \sum_{i=1}^{n} t_i c_i r_i}{\mathrm{vc} + \sum_{i=1}^{n} t_i c_i} .
$$

with c_i being the contribution of review r_i to p_d. If r_i is a direct review on p_d, $c_i = 1$. If r_i is a review of a document p_j, c_i is the sum of the contributions of r_i along all paths $[\ p_j \to \ldots \to p_d\ ]$ with path length less or equal to $k_{\max}$.

3.4 Distance-based TRE-Visibility

We can simplify the tre-visibility calculation by determining the contribution of a review r_i solely by its distance $k_{i,d}$ to p_d, following the idea that a review should have a larger impact if it is closer to the paper.[15] With β for finetuning the impact of indirect reviews, we get:

$$
\mathrm{vis}_{p_d} = \frac{\mathrm{vc} \cdot \mathrm{vis}^{\circ}_{p_d} + \sum_{i=1}^{n} \left(\frac{t_i}{(k_{i,d}+1)^{\beta}} \cdot r_i \right)}{\mathrm{vc} + \sum_{i=1}^{n} \frac{t_i}{(k_{i,d}+1)^{\beta}}}
$$

---

[12] Obviously this can be done in a similar way for other visibility functions.

[13] This also holds for other recursive approaches for integrating a second, user-dependent source of information in a recursive visibility function, see e. g. the twr-visibility function of [Hess *et al.*, 2006].

[14] This can be done very efficiently by simply propagating all reviews of a document p_i to all documents $p_a \in C_i$ cited by it, annotated with the outdegree $|C_i|$ and repeated for $k_{\max}$ steps.

[15] This is implicitly also true for the tre$_p$-visibility if each document in a path cites more than one other document, and the average number of documents cited is larger than the number of paths.

3.5 Efficient Computation

Discussing the runtime of the functions described above, we have to distinguish the overall costs and costs at query time. While offline computation of large document repositories is costly, it is feasible as e. g. Google shows. Critical is the time needed to answer a single search query. As mentioned before it is impossible to compute the visibilities of all documents of a large document repository at query time. On the other hand, it is also impossible to precompute personalized rankings of all documents for all users, as this would even offline be to costly. Thus tre_i is at least not appropriate for personalized search engines.

We could reduce computation costs by restricting the visibility computation to the subset of documents we are interested in (e. g. the set of documents matching the search term). Unfortunately, this is not possible for tre_i because the tre_i-visibility of the documents in the subset depends on the tre_i-visibility of the documents they are cited by and so on.[16]

Here the other tre-visibility functions come into play: the data on both subnets can be precomputed, and only the join has to be done on the fly. The trust of any user in any other user can be precomputed offline, which automatically gives the trust in each review. Re-computation is required if trust edges change or new users join (but not, if new reviews are given!). The base visibility $\text{vis}^{\circ}_{p_j}$ of each document p_j can also be precomputed. The simple tre-visibility tre_s of the subset of documents the user is interested in can now be computed in a single run through these documents. For using tre_p or tre_d, each document additionally has to know its indirect reviews, and these sets of reviews can also be precomputed by simply propagating all reviews through the document network while computing the base document visibility. New reviews are propagated through the document network at the time they are created. So at query time all to be done is the join: for all documents p_j in the selected subset, all direct and indirect reviews and the document base visibility $\text{vis}^{\circ}_{p_j}$ are summed up, weighted by their corresponding trust values in order to compute $\text{vis}^{\text{tre}}_{p_j}$. And finally the documents are sorted to provide a personalized ranking.

4 Simulation

It is obvious that the functions presented take direct (and indirect) reviews weighted by their trust values into account and compute personalized tre-visibilities. The interesting question is how much the presented functions differ and which impact they give to indirect reviews.

We compare the different tre-visibility functions by computing tre-visibilities from the perspective of a test user u on 10 document reference networks, each with ≈ 12000 documents (with 2 to 7 references) and 1000 reviews randomly distributed. The test user's trust in each review was uniformly distributed in $[0,1]$. The document base visibilities were computed by PageRank with $\alpha = 0.85$ and $N = 100$. Now the tre-visibilities of all documents were computed using tre_s, tre_i, tre_p and tre_d (with vc $= 0.5$, $k_{\max} = 3$, $\beta = 3$). By comparing the visibilities ($\text{vis}^A_{p_d}$, $\text{vis}^B_{p_d}$) of a node p_d as computed by two visibility algorithms A and B we can see how big the differences between A and B are. If A and B would compute the same function, the difference would be 0. Table 1 gives the average difference of the visibilities computed by algorithm A and B, respectively, over all documents: Δ_{direct} is the average difference of documents with at least one direct review, Δ_{indirect} the same for documents without direct review and Δ_{total} the overall difference.[17]

Alg. A	Alg. B	Δ_{direct}	Δ_{indirect}	Δ_{total}
PageRank	tre_s	0.228	0	0.019
PageRank	tre_i	0.267	0.075	0.091
PageRank	tre_d	0.256	0.077	0.092
PageRank	tre_p	0.257	0.079	0.094
tre_s	tre_i	0.040	0.075	0.072
tre_s	tre_d	0.030	0.077	0.073
tre_s	tre_p	0.031	0.079	0.075
tre_i	tre_d	0.024	0.043	0.042
tre_i	tre_p	0.025	0.046	0.044
tre_d	tre_p	0.010	0.020	0.019

Table 1: Differences in visibility computation

The first four lines of Tab. 1 show the differences to PageRank. Obviously, direct reviews have high impact on all tre-visibility functions, but also the effect of indirect reviews is considerable (with the exception of tre_s, certainly). In the next three lines Δ_{direct} is most interesting, which shows to which amount the visibility of directly reviewed documents is changed by additional indirect reviews (which are not considered by tre_s but by the other functions). The next two lines show (Δ_{indirect}) that tre_d and tre_p differ from tre_i (we did not expect them to totally resemble tre_i), but much less than from PageRank. And finally the last line shows, that tre_p and tre_d are very close so that for the given networks the average number of 4.5 outgoing citations per document (affecting tre_p) is well resembled by setting $\beta = 3$ (pretests with $\beta = 2$ gave large differences). This also shows that the parameters of the functions used have to be adjusted for each network size and structure to give appropriate results.

5 Conclusion

In this paper, we introduced different trust-enhanced visibility function integrating document visibilities, user-dependent trust information and reviews for personalized document recommendations. We attached importance to design functions that calculate personalized rankings efficiently at query time.

[16]This does not fully hold for HITS, where the computation is in fact done on a subset.

[17]We did this separately for all 10 networks, but the differences between them were so small (standard deviation $s < 1\%$), that we only show the average over all networks.

Therefore, we analyzed how to use pre-calculated information on both networks that is joint on the fly. The simulation study compared the functions, and we can conclude that the path-based and the distance-based tre-visibility permit computing personalized recommendations of the same quality as the costly recursive tre-visibility while being efficiently computable, and thus appropriate for search engines. On the basis of these results, we now aim to develop a personalized recommender system integrating our efficient trust-enhanced visibility computation.

Acknowledgments

This work was partially supported by the German Academic Exchange Service (DAAD) and the German Science Foundation (DFG).

References

[Avesani *et al.*, 2005] Paolo Avesani, Paolo Massa, and Roberto Tiella. A trust-enhanced recommender system application: Moleskiing. In *Proceedings of the 20th Annual ACM Symposium on Applied Computing*, pages 1589–1593, 2005.

[Budd *et al.*, 1998] John M. Budd, MaryEllen Sievert, and Tom R. Schultz. Phenomena of retraction: Reasons for retraction and citations to the publications. *JAMA*, 280(3):296–297, July 1998.

[Dumbill, 2002] Edd Dumbill. Finding friends with xml and rdf. IBM's XML Watch, June 2002.

[García-Barriocanal and Sicilia, 2005] Elena García-Barriocanal and Miguel-Angel Sicilia. Filtering information with imprecise social criteria: A foaf-based backlink model. In *4th Conference of the European Society for Fuzzy Logic and Technology EUSFLAT*, Barcelona, Spain, 2005.

[Golbeck *et al.*, 2003] Jennifer Golbeck, Bjian Parsia, and James Hendler. Trust networks on the semantic web. In *Proceedings of Cooperative Intelligent Agents*, Helsinki, Finland, August 2003.

[Golbeck, 2006] Jennifer Golbeck. Generating predictive movie recommendations from trust in social networks. In *Proceedings of the Fourth International Conference on Trust Management*, Pisa, Italy, May 2006.

[Gyöngyi *et al.*, 2004] Zoltán Gyöngyi, Hector Garcia-Molina, and Jan Pedersen. Combating web spam with trustrank. In *Proceedings of the 30th International Conference on Very Large Data Bases (VLDB)*, pages 271–279, 2004.

[Hess *et al.*, 2006] Claudia Hess, Klaus Stein, and Christoph Schlieder. Trust-enhanced visibility for personalized document recommendations. In *Proceedings of the 21st Annual ACM Symposium on Applied Computing*, Dijon, France, April 2006.

[Kleinberg, 1999] Jon M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5):604–632, September 1999.

[Korfiatis and Naeve, 2005] Nikolaos Korfiatis and Ambjörn Naeve. Evaluating wiki contributions using social networks: A case study on wikipedia. In *First on-Line conference on Metadata and Semantics Research (MTSR'05)*, 2005.

[Massa and Avesani, 2004] Paolo Massa and Paolo Avesani. Trust-aware collaborative filtering for recommender systems. In *Proceedings of the International Conference on Cooperative Information Systems (CoopIS)*, 2004.

[Page *et al.*, 1998] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford Digital Library Technologies Project, 1998.

[Pinski and Narin, 1976] Gabriel Pinski and Francis Narin. Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics. *Information Processing & Management*, 12:297–312, 1976.

[Stein and Hess, 2006] Klaus Stein and Claudia Hess. Information retrieval in trust-enhanced document networks. In M. Ackermann, B. Berendt, M. Grobelnik, A. Hotho, D. Mladenic, G. Semeraro, M. Spiliopoulou, G. Stumme, V. Svatek, and M. van Someren, editors, *Semantics, Web, and Mining. European Web Mining Forum, EMWF 2005, and Knowledge Discovery and Ontologies, KDO 2005, Porto, Portugal, October 2005, Revised Selected and Invited Papers*, LNAI 4289. Springer, 2006.