

Time-Aware Authority Ranking

Klaus Berberich, Michalis Vazirgiannis, and Gerhard Weikum

Abstract. The link structure of the web is analyzed to measure the authority of pages, which can be taken into account for ranking query results. Due to the enormous dynamics of the web, with millions of pages created, updated, deleted, and linked to every day, temporal aspects of web pages and links are crucial factors for their evaluation. Users are interested in important pages (i.e., pages with high authority score) but are equally interested in the recency of information. Time—and thus the freshness of web content and link structure—emanates as a factor that should be taken into account in link analysis when computing the importance of a page. So far only minor effort has been spent on the integration of temporal aspects into link-analysis techniques. In this paper we introduce T-Rank Light and T-Rank, two link-analysis approaches that take into account the temporal aspects freshness (i.e., timestamps of most recent updates) and activity (i.e., update rates) of pages and links. Experimental results show that T-Rank Light and T-Rank can produce better rankings of web pages.

I. Introduction

The web graph grows at a tremendous rate while its content is updated at a very high pace following interesting patterns [Fetterly et al. 03]. The issue of analyzing the link structure of the web in order to determine a page's authority has attracted significant efforts [Kleinberg 99, Page et al. 98].

The predominant authority ranking approach in search engines is the Page-Rank method that was proposed by Page et al. [Page et al. 98]. However, this is a precomputed, static measure that does not reflect the temporal aspects accompanying the evolution of pages and links. Due to the high dynamics of the

web, with millions of pages created, updated, deleted, and linked to every day, temporal aspects of web pages and links are crucial factors for their evaluation. Users are interested not just in the most important pages (i.e., the ones with the highest authority) but in information that is both important and recent. For example, when we submitted the query “VLDB Conference” in September 2004, none of the major search engines returned the web pages belonging to the VLDB Conferences 2004 and 2005 among its five highest ranked authorities, although it is likely that a user submitting this query at that specific time was interested in one of those two web pages. Thus, as the example illustrates, the users’ interest has a temporal dimension, which could be, for example, on the present time. This temporal interest, as we call it, should be considered for the authority ranking of web pages.

We consider two temporal aspects of the web’s evolution in this paper: the time when a web page or link was last updated, which is referred to as *freshness*, and, as a second temporal aspect, the rate of updates of a page’s content and its incoming links (in-links), referred to as *activity*. It is intuitive that pages whose content or in-links were frequently updated with regard to a user’s temporal interest are more interesting to the user. On the contrary, pages whose content and in-links were not updated regarding the user’s temporal interest indicate lack of activity and therefore minor interest from the web community.

The seminal works of Page et al. [Brin and Page 98, Page et al. 98] and Kleinberg [Kleinberg 99] were followed by rich work on link analysis and authority ranking for the web [Amitay et al. 04, Baeza-Yates et al. 02, Haveliwala 02, Jeh and Widom 03, Lempel and Moran 01, Xue et al. 03]. However, only minor effort has been spent on the integration of temporal aspects into ranking techniques for web or intranet searches. Amitay et al. [Amitay et al. 04] use the last modification time of a web page to weight its outgoing hyperlinks and on this basis apply a combination of the HITS [Kleinberg 99] and SALSA [Lempel and Moran 01] approaches to assess authority. Another related effort by Baeza-Yates et al. [Baeza-Yates et al. 02] observes that PageRank favors old pages. To counter this bias, the authors devise variants of PageRank that take into account the age of web pages. In their recent work, Yu et al. [Yu et al. 04] propose another variant of PageRank that they apply to a network of scientific publications. In their approach, citations are weighted based on their age, and a post-processing step decays the authority of a publication based on the publication’s age. Bar-Yossef et al. [Bar-Yossef et al. 04], finally, propose to assess the decay of a web page by means of the link structure. Here, a web page is regarded more outdated if it (transitively) points to web pages that have disappeared.

The web graph and its evolution have been the subject of several recent studies [Broder et al. 00, Fetterly et al. 03, Kumar et al. 00, Ntoulas et al. 04]. Models

for describing the evolving web graph were proposed by Kraft et al. [Kraft et al. 03] and Kumar et al. [Kumar et al. 05]. The former group of authors extracts parts of the evolving web graph from the Internet Archive [Archive 05] and presents statistics on the extracted data. The latter group of authors analyzed the dynamics of a well-defined subspace of the web, namely a collection of so-called weblogs.

The time-aware methods T-Rank Light and T-Rank developed in this paper extend link analysis by taking into account the temporal aspects freshness and activity of both pages and links. Technically, our approach can be seen as an extension of the PageRank-style family of random-walk models. Just like PageRank we compute stationary state probabilities of pages for an appropriately defined Markov chain by numerical power iteration.

We experimentally evaluated the proposed time-aware methods. The results of these experiments indicate that the methods can indeed significantly improve the rankings. Our techniques are not limited to the web but can also be applied to any kind of graph structure that bears appropriate temporal information. Thus, in a first experiment, we apply the methods to the Digital Bibliography & Library Project (DBLP) authors' bibliographic network. In a second experiment, the methods are evaluated on a graph structure of product information that is obtained from Amazon.com. Finally, in a third experiment, we evaluate the methods based on web data collected before and during the 2004 Olympic Games.

2. Model

The building blocks of the formal model are the description of the evolving graph and the specification of the user's temporal interest.

2.1. Evolving Graph

The directed graph $G(V, E)$ consists of the set of nodes V and the set of edges E . Both nodes and edges in this graph are annotated with temporal information. In our model, time is represented by integers, implying that time is discretized at a fixed granularity. Timestamps correspond to the number of time units, according to the chosen granularity, that have passed since a reference time. The temporal information annotated to nodes and edges in the graph indicates different kinds of events. For each node and edge, one timestamp $TS_{Creation}$ maintains when the node or edge was created; similarly, a timestamp $TS_{Deletion}$ indicates when the node or edge was deleted. Thus, for every node and edge, its lifespan is captured by the interval $[TS_{Creation}, TS_{Deletion}]$. In a concrete application

scenario (e.g., web search) nodes and edges may be deleted but reappear at a later time. Here, a possible mapping of the real world onto our model could, for instance, define the interval $[TS_{Creation}, TS_{Deletion}]$ for a node or edge to cover the time between its earliest and last observation. Moreover, the model captures modifications of nodes and edges, hence each node and edge is annotated with a set of timestamps $TS_{Modifications}$, members of which correspond to the times when the node or edge was modified. These modifications do not necessarily correspond to structural changes of the graph: if we use the proposed model, e.g., for describing the web's evolution, we would map a modification of a web page to a modification of the corresponding node and a modification of a hyperlink's anchor text to a modification of the corresponding edge. We demand the following constraints for the temporal annotations on nodes and edges. First of all, a node or an edge must be created before it is deleted, i.e.,

$$TS_{Creation} \leq TS_{Deletion}. \quad (2.1)$$

Moreover, we demand that all modifications to a node or an edge are applied during its lifespan, that is to say

$$\forall m \in TS_{Modifications} : m \in [TS_{Creation}, TS_{Deletion}]. \quad (2.2)$$

2.2. Temporal Interest

A further part of the model is the user's temporal interest, which represents the temporal window in which the user is interested. The temporal interest consists of two temporal intervals that we call the *temporal window of interest* and the *tolerance interval*. The temporal window of interest is represented by two timestamps TS_{Origin} and TS_{End} that fulfill $TS_{Origin} \leq TS_{End}$ and specify the temporal interval in which the user is interested. The tolerance interval is defined by two timestamps t_1 and t_2 . It extends the temporal window of interest, that is to say timestamps inside the tolerance interval but outside the temporal window of interest are assumed to be less interesting to the user. Furthermore, we demand for the limits of the two temporal intervals that $t_1 \leq TS_{Origin} \leq TS_{End} \leq t_2$ holds. Hence, the tolerance interval $[t_1, t_2]$ can surround the temporal window of interest $[TS_{Origin}, TS_{End}]$, but also the case where one or both limits coincide with the limits of the temporal window of interest is possible. Moreover, the limits of both temporal intervals can coincide with a single timestamp, thus indicating that the users is only interested in this single time.

2.3. Freshness and Activity

Based on the evolving graph described in Section 2.1 and the temporal interest, we formally define freshness and activity as the two temporal aspects taken into

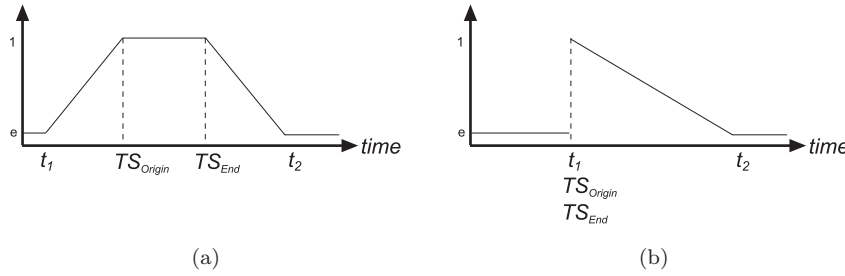


Figure 1. Temporal interest examples: (a) the Olympic Games and (b) a terrorist attack.

account by the proposed time-aware methods. As stated in the introduction, freshness reflects whether a node or an edge, i.e., a page or a hyperlink, is up to date with regard to the user's temporal interest. We define freshness as a function f . Given a temporal interest and a timestamp ts , freshness is maximal if ts is inside the temporal window of interest (i.e., $ts \in [TS_{Origin}, TS_{End}]$). Freshness decreases linearly with the distance to the temporal window of interest, if ts is only in the tolerance interval but not inside the temporal window of interest (i.e., $ts \in [t_1, t_2] \setminus [TS_{Origin}, TS_{End}]$). Otherwise, if ts is outside the tolerance interval, a minimal freshness value $e > 0$ is assumed, thus avoiding that freshness converges to 0. The complete formal definition of freshness is as follows:

$$f(ts) = \begin{cases} \frac{1-e}{TS_{Origin}-t_1} \cdot (ts - t_1) + e & : t_1 \leq ts < TS_{Origin} \\ 1 & : TS_{Origin} \leq ts \leq TS_{End} \\ \frac{e-1}{t_2-TS_{End}} \cdot (ts - TS_{End}) + 1 & : TS_{End} < ts \leq t_2 \\ e & : \text{otherwise.} \end{cases}$$

As an illustration, we give two temporal interests and depict the associated freshness functions. Assume, for instance, a user interested mainly in an event with a fixed beginning and end like the Olympic Games, as well as in the prior and following discussion (see Figure 1(a)). In this case, we set the temporal window of interest to include the duration of the event (i.e., $TS_{Origin} = \text{August 15, 2004}$, and $TS_{End} = \text{August 31, 2004}$). The tolerance interval is set to include, for example, the period one month before and after the event (i.e., $t_1 = \text{July 15, 2004}$, and $t_2 = \text{September 30, 2004}$) to consider the discussions, which are assumed to be of lower interest to the user. We can also handle the case where a user is interested in the period after an unexpected event, e.g., the unfortunate terrorist attack in Madrid in March 2004 (see Figure 1(b)). Here, there is no need to

include times before the event itself, so we set $t_1 = TS_{Origin} = TS_{End}$ and, in order to include things in the aftermath with decreasing freshness according to their distance, we set $t_2 > t_1$.

Freshness is adapted to objects, i.e., nodes and edges, in the following way. Given an object o and its annotated temporal information, we define its freshness as the maximal freshness of its creation time and modification times, i.e.,

$$f(o) = \text{Max}\{f(ts) \mid ts \in TS_{Modifications}(o) \cup \{TS_{Creation}(o)\}\}.$$

Moreover, including the creation time $TS_{Creation}$, we guarantee that the freshness of an object is always defined and nonzero.

Activity is the second temporal aspect to be formally defined. As stated before, activity conveys the frequency of change. The basis for computing the activity of an object is the annotated set $TS_{Modifications}$. We define the activity of an object o as the sum of the freshness values of modification timestamps within the tolerance interval, i.e., the timestamps in $TS_{Modifications} \cap [t_1, t_2]$. Formally, the function a is defined as follows:

$$a(o) = \sum f(ts) \quad \text{with} \quad ts \in (TS_{Modifications}(o) \cap [t_1, t_2]) \cup \{TS_{Creation}(o)\}.$$

Note that, by definition, activity takes into account the time when an object was modified. If we assume, for instance, two objects A and B that were modified an equal number of times but A has modifications at times of higher freshness, it is also assigned higher activity, since it was modified at times more interesting to the user. However, the freshness value of object B could be larger if the maximal freshness of B 's modification times is larger than the corresponding value of A . Moreover, the creation time $TS_{Creation}$ is included in the definition of activity, regardless of when the object was created, thus guaranteeing that the activity value of an object is always nonzero.

2.4. Graph with Respect to the Temporal Interest

In addition to the evolving graph, the temporal interest, and the temporal aspects defined in this section, we define G_{ti} as the *graph with respect to the temporal interest*. Here, the idea is to prune the evolving graph based on a given temporal interest to eliminate nodes and edges that do not exist at any time, which is of interest to the user. Similar concepts to prune the graph with regard to a point in time or a time interval were included in both aforementioned graph models that were proposed by Kraft et al. [Kraft et al. 03] and Kumar et al. [Kumar et al. 05], respectively.

G_{ti} consists of the set of nodes V_{ti} and E_{ti} , which are subsets of the corresponding constituents of the evolving graph. Formally, the set of nodes V_{ti} is

defined as follows:

$$V_{ti} = \{v \in V \mid TS_{Creation}(v) \leq t_2 \wedge TS_{Deletion}(v) \geq t_1\}.$$

Thus, only nodes with a lifespan $[TS_{Creation}, TS_{Deletion}]$ that overlaps the tolerance interval are included in G_{ti} . Similarly, we define the set of edges E_{ti} as follows:

$$E_{ti} = \{(x, y) \in E \mid (x, y) \in V_{ti} \times V_{ti} \wedge TS_{Creation}(x, y) \leq t_2 \wedge TS_{Deletion}(x, y) \geq t_1\}.$$

By this definition, only edges having a lifespan $[TS_{Creation}, TS_{Deletion}]$ that overlaps the tolerance interval, i.e., edges that existed at any point of it, are included in G_{ti} .

Consider that, for a temporal interest whose temporal window of interest and tolerance interval coincide with a single timestamp, the graph G_{ti} gives a snapshot of the graph at this specific time. In the remainder of this work, we apply our time-aware methods and PageRank, as their “competitor,” on the basis of the graph G_{ti} for a given temporal interest.

3. Time-Aware Authority Ranking

The time-aware methods build on and extend the PageRank method [Brin and Page 98, Page et al. 98], which is based on a random-walk model.

3.1. PageRank

In the PageRank method a random surfer travels the (web) graph by either following outgoing links or by randomly jumping to another node. In both cases the next page to be visited is chosen according to a uniform probability distribution among the possible target pages. The event of making a random jump occurs with a fixed probability ϵ . The PageRank measure of a node in the graph is its stationary state probability, or equivalently, the fraction of time the random surfer spends on this page as the length of the random walk approaches infinity. The following formula gives the PageRank $r(y)$ of a node y :

$$r(y) = (1 - \epsilon) \left(\sum_{(x,y) \in E} \frac{r(x)}{\text{outdegree}(x)} \right) + \frac{\epsilon}{n}. \quad (3.1)$$

An interpretation of this recursive equation is that every page “distributes” its authority uniformly to all its successors, which is captured in the first term of the

formula. This corresponds to the random surfer following an outgoing edge with probability ϵ and then choosing a specific outgoing edge with uniform probability $1/\text{outdegree}(x)$. The second term in the equation corresponds to the authority that a node gains when it is the target of a random jump.

The original PageRank work [Page et al. 98] described already the possibility of modifying the probability distribution underlying the random jump to favor certain nodes. The intention was to allow, for example, personalized or topic-sensitive authority scores. This idea has been adopted and developed further by Jeh and Widom [Jeh and Widom 03] and Haveliwala [Haveliwala 02]. The approaches proposed by Xue et al. [Xue et al. 03] and Bazea-Yates et al. [Bazea-Yates et al. 02] further propose to use nonuniform probability distributions for the transition probabilities. Our generalization of the PageRank method, which is described in the following formula, introduces a bias for both transition probabilities and random jump probabilities:

$$r(y) = (1 - \epsilon) \left(\sum_{(x,y) \in E} t(x,y) \cdot r(x) \right) + \epsilon \cdot s(y) \quad (3.2)$$

$$\text{with } \sum_y s(y) = 1 \quad \text{and} \quad \sum_y t(x,y) = 1$$

Here the function t describes the transition probabilities. A specific value $t(x,y)$ corresponds to the probability that the random surfer, being in node x and having decided to make a transition, chooses to go to node y . The function s describes the random jump probabilities, a probability distribution which is independent of the random surfer's current position.

Employing this framework, the standard PageRank method, which is applied to the graph G_{ti} and serves as a "competitor" to the time-aware methods, can be described as follows. The uniform random jump probabilities are described by a function s , so that the function value $s(y)$ is defined as

$$s(y) = |V_{ti}|^{-1}. \quad (3.3)$$

The uniform transition probabilities are defined as a function t , so that the function value $t(x,y)$ assumes 0 if no edge exists between the nodes x and y , i.e., $(x,y) \notin E_{ti}$. Otherwise, we define

$$t(x,y) = |\{(x,z) \mid (x,z) \in E_{ti}\}|^{-1}. \quad (3.4)$$

Casting Equation (3.2) into matrix form and applying some simple transformations lead to

$$\pi = M^T \cdot \pi. \quad (3.5)$$

Here, the entries of the column vector π , which is the principal eigenvector of the stochastic matrix M , correspond to the PageRank scores of the nodes in the graph. The matrix M is defined as $M = (1-\epsilon)\cdot T + \epsilon\cdot S$, with matrices $T_{ij} = t(i, j)$ and $S_{ij} = s(j)$. Thus, the matrix T describes the transition probabilities, and the matrix S captures the random jump probabilities. The matrix M , finally, is the transition probability matrix of the Markov chain defined by the random surfer's random walk. Consider that, for the matrix M to be stochastic, the graph is not allowed to have dangling nodes. However, if dangling nodes exist, we can make the matrix M stochastic by setting $t(x, y) = s(y)$ for a dangling node x . This fix corresponds to the random surfer making a random jump whenever no outgoing edge is available for transition. Solving Equation (3.5) for the vector π is a principal eigenvector computation, which can be numerically handled using the power iteration method. Thus, we choose an initial probability vector $x^{(0)}$ (e.g., a uniform vector with $x_i^{(0)} = 1/n$) and compute the result of the $(k+1)$ th iteration by applying

$$x^{(k+1)} = M^T \cdot x^{(k)}. \quad (3.6)$$

The computation is stopped when the size of the residual vector is smaller than a threshold value δ , i.e.,

$$\|x^{(k+1)} - x^{(k)}\|_1 < \delta. \quad (3.7)$$

3.2. T-Rank Light

T-Rank Light, as the first time-aware method proposed in this paper, adopts the idea by Page et al. [Page et al. 98] to skew the random jump probabilities to favor certain nodes in the graph. More precisely, the random jump probabilities are affected by the freshness and the activity of nodes and further by the average freshness and activity of the incoming edges to a node.

On the web the freshness of a node, on the one hand, conveys whether the corresponding web page was modified with regard to the user's temporal interest. The activity of a node, on the other hand, conveys the frequency of change of the corresponding web page regarding the temporal interest. Thus, on the web, both temporal aspects depend on the associated web page and are therefore susceptible to manipulation, since a malicious site owner can accomplish high freshness and activity by applying frequent modifications. On the contrary, the average temporal information on incoming edges to a node is normally not under the control of a single site owner but depends on a multitude of web pages. In detail, the average freshness and activity of incoming edges to a node reflect the attention the corresponding web page has attracted regarding the temporal interest.

In comparison to PageRank, the random walk underlying T-Rank Light is modified as follows: when making a random jump, the random surfer selects the target of the random jump based on the temporal information delineated in Section 2.3. More precisely, first one of the four different kinds of temporal information is randomly chosen with fixed probabilities. In a second step the target of the random jump is randomly selected with probabilities proportional to the previously chosen kind of temporal information. Apart from that, the random walk is not modified, i.e., for the transitions the random surfer behaves as in PageRank.

For the formal definition we introduce coefficients w_{si} that give the probabilities of the random surfer selecting one of the four kinds of temporal information to choose the target of the random jump. The coefficients must describe a valid probability distribution, i.e., we demand $w_{si} \geq 0$ and $\sum w_{si} = 1$. Applying appropriate normalization, this leads us to the following formal definition of the function s describing the random jump probabilities:

$$\begin{aligned}
 s(y) = w_{s1} \cdot \frac{f(y)}{\sum_{z \in V_{ti}} f(z)} \\
 + w_{s2} \cdot \frac{\text{avg}\{f(x, y) \mid (x, y) \in E_{ti}\}}{\sum_{z \in V_{ti}} \text{avg}\{f(x, z) \mid (x, z) \in E_{ti}\}} \\
 + w_{s3} \cdot \frac{a(y)}{\sum_{z \in V_{ti}} a(z)} \\
 + w_{s4} \cdot \frac{\text{avg}\{a(x, y) \mid (x, y) \in E_{ti}\}}{\sum_{z \in V_{ti}} \text{avg}\{a(x, z) \mid (x, z) \in E_{ti}\}}.
 \end{aligned} \tag{3.8}$$

The formula given in Equation (3.8) captures the random surfer's behavior as described above. The first summand, for instance, has the following intuition. In the first step, when choosing the criterion based on which the random jump is performed, the random surfer chooses the *freshness of nodes* with probability w_{s1} and then, in the second step, chooses the node y (among all nodes in the graph) with probability proportional to its freshness $f(y)$. The denominator of the fraction simply sums up the freshness of all nodes for normalization. The structure of the other summands has analogous intuitions.

By virtue of the definition of freshness and activity, the denominators in Equation (3.8) are guaranteed not to assume zero values. The function avg incorporates the average, which is the mean in our concrete implementation but could be replaced by another average or aggregation function. As stated earlier, in case of a transition, the random surfer adopts the behavior described for PageRank. Consequently, the function t describing the transition probabilities is the same that was given in Equation (3.4). For the computation of the station-

ary state probabilities of the Markov chain underlying the random walk defined by T-Rank Light, the power iteration method as introduced in Section 3.1 can be used.

For the Markov chain defined by standard PageRank, stationary state probabilities are guaranteed to exist. The uniform random jump probabilities are crucial for this guarantee, since they assure that the Markov chain is irreducible, i.e., every pair of states is mutually reachable. For the function s defined in Section 3.1, irreducibility is only guaranteed for parameters fulfilling $w_{s1} + w_{s3} > 0$. However, other work [Haveliwala 02] reported that no problems occur in practice even if arbitrary random jump probabilities are used. On the basis of our experiments, we second this statement, since we did not face any problems using arbitrary random jump probabilities.

3.3. T-Rank

For T-Rank, our second time-aware method, we additionally skew the transition probabilities based on temporal information. Figure 2 illustrates the situation of a random surfer being in node x and making a transition. In the figure the label t indicates that freshness and activity values are available for the labeled node or edge. As can be seen from the figure, the random surfer is aware of the temporal information that was used for the definition of the random jump, i.e., the freshness and activity of nodes and the freshness and activity of incoming edges to a node. In addition, the random surfer is aware of the temporal information on those edges connecting the random surfer's current position (i.e., the node x in the figure) and the possible next positions (i.e., the nodes y and z). Thus, we obtain the freshness and activity of these interconnecting edges as two further kinds of temporal information taken into account for the definition of the transition probabilities. On the web, the semantics associated with these two kinds of temporal information is as follows: if an edge between the random surfer's current position and a successor was created or modified with regard to the temporal interest, this conveys that the web page associated with the current node paid attention to the successor with regard to the temporal interest.

The random walk underlying T-Rank adopts the basic structure from PageRank, and thus we use the generalized PageRank as a means of description. Regarding the random jump probabilities, we see that the random surfer's be-

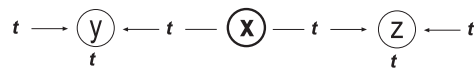


Figure 2. Random surfer in node x .

havior is exactly the same as that defined for T-Rank Light and described in the previous section. In a similar manner we define the random surfer's behavior in case of a transition. Thus, when making a transition, the random surfer first selects one of the six kinds of temporal information identified in the previous paragraph with fixed probabilities and then randomly chooses one of the successors of the current position with probabilities proportional to the selected kind of temporal information.

For the formal definition of the random jump probabilities, we refer to Equation (3.8) that defined the random jump probabilities as a function s . For the definition of the transition probabilities, we introduce coefficients w_{ti} , which give the probabilities of the random surfer selecting the next position based on the six different kinds of temporal information. We demand that these coefficients describe a probability distribution, i.e., $w_{ti} \geq 0$ and $\sum w_{ti} = 1$. The transition probabilities are defined as a function t , so that $t(x, y)$ assumes 0 if no edge exists between the nodes x and y , i.e., $(x, y) \notin E_{ti}$. Otherwise, we define $t(x, y)$ as the following weighted sum:

$$\begin{aligned}
 t(x, y) = & w_{t1} \cdot \frac{f(y)}{\sum_{(x,z) \in E_{t1}} f(z)} \\
 & + w_{t2} \cdot \frac{f(x, y)}{\sum_{(x,z) \in E_{t2}} f(x, z)} \\
 & + w_{t3} \cdot \frac{\text{avg}\{f(v, y) \mid (v, y) \in E_{t3}\}}{\sum_{(x,z) \in E_{t3}} \text{avg}\{f(w, z) \mid (w, z) \in E_{t3}\}} \\
 & + w_{t4} \cdot \frac{a(y)}{\sum_{(x,z) \in E_{t4}} a(z)} \\
 & + w_{t5} \cdot \frac{a(x, y)}{\sum_{(x,z) \in E_{t5}} a(x, z)} \\
 & + w_{t6} \cdot \frac{\text{avg}\{a(v, y) \mid (v, y) \in E_{t6}\}}{\sum_{(x,z) \in E_{t6}} \text{avg}\{a(w, z) \mid (w, z) \in E_{t6}\}}.
 \end{aligned} \tag{3.9}$$

Equation (3.9) captures the behavior of the random surfer as described above. Intuitively, the third summand, for instance, states that with probability w_{t3} in a first step the random surfer chooses the next node based on the average freshness of incoming edges. In a second step, the next node is selected among the successors of node x with probability proportional to the *average freshness of incoming edges* of the successors.

The previously-defined functions s and t and the generalized PageRank yield a full formal definition of T-Rank. The authority scores according to T-Rank,

i.e., the stationary state probabilities of the underlying Markov chain, can be computed by applying the power iteration method as described in Section 3.1. Again, the structure of the other summands has analogous intuitions.

T-Rank Light and T-Rank are defined incrementally, thus varying in the extent of bias introduced to the uniform random walk defined by PageRank. Regarding the applicability and scalability of the two methods, we make the following observation: variants of PageRank that rely on a skewing of the random jump vector, and are therefore technically related to T-Rank Light, have been applied to relatively large datasets of up to 24 million pages, e.g., by Page et al. [Page et al. 98] and Haveliwala [Haveliwala 02]. More generally, if we assume, for instance, a dataset containing a billion web pages, the storage cost of storing the function values $s(y)$ at double-precision (we assume 64 bit values here) without applying any compression is about eight gigabytes. The application of the power iteration method only requires linearly scanning these eight gigabytes at the end of each iteration. When facing a shortage of main memory, we can store the $s(y)$ values on disk and fetch them into memory in large chunks as needed by the T-Rank Light computation. With sequential disk access and asynchronous prefetching, this does not slow down the computation while requiring only a modest amount of memory (e.g., in the order of a megabyte, equivalent to several disk tracks). Note that 64 bits per score value is very generous; this could easily be compressed by an order of magnitude. Therefore, T-Rank Light is applicable to very large datasets. We are not aware of any prior variant of PageRank that also skews the transition probabilities and is applied to comparably large datasets. Baeza-Yates et al. [Baeza-Yates et al. 02, Baeza-Yates et al. 04] apply their methods to a dataset of 795,000 pages, and Xue et al. [Xue et al. 03] apply their methods to the moderately-sized graphs to be found in an intranet, which has only 170,000 pages. Applying such methods, or T-Rank, to web-scale data volumes poses a challenge. In contrast, T-Rank Light does not face such scalability limitations.

4. Experiments

We implemented PageRank, T-Rank Light, and T-Rank in a prototype system using the Java programming language. Our implementation is application-independent in the sense that it is applicable to any dataset that can be mapped to an evolving graph as described in Section 2.1. Our experimental evaluation of the time-aware methods comprises of experiments on a bibliographic dataset derived from DBLP, experiments on a dataset of e-commerce product data obtained from Amazon.com, and a web dataset collected before and during the

2004 Olympic Games in Athens. In the following, we present the results of our experiments on the different datasets and, following that, discuss our findings on the parameter sensitivity of the methods.

4.1. Experiment on DBLP

The Digital Bibliography and Library Project (DBLP) provides bibliographic information on more than half a million computer science publications. Unfortunately, citation information is only recorded for a subset of the publications, mostly being related to the field of databases. We derived a graph structure from this data as follows: authors are mapped to nodes in the graph; citations between (papers of) authors map to directed edges. The temporal annotation $TS_{Creation}$ of a node corresponds to the date of the corresponding author's first publication. For the subsequent publications of the author, entries to the annotated set $TS_{Modifications}$ are created. Similarly, an edge (x, y) is assumed to be created when the author corresponding to node x cited the author corresponding to node y for the first time. Whenever the author corresponding to node x cites the author corresponding to node y repeatedly an entry to the set $TS_{Modifications}$ is created. In total, the obtained graph structure has 347,986 nodes and 346,797 edges. Since citations are only scarcely recorded and span mostly the field of databases, only 16,316 nodes in the graph are connected, i.e., have either incoming or outgoing edges.

The recorded publications span the period from the 1970s to now with the number of publications per year increasing almost steadily over time. Regarding the recorded citations, we observe that most citations are recorded for the 1980s and 1990s, whereas there is only little citation information for before the 1980s and after the new millennium. Moreover, we investigated the distribution of in-degrees and out-degrees in the obtained graph structure. For the in-degrees we observed that they fit pretty well a power law distribution of the form $P[(in|out)degree = k] \sim (1/k)^\alpha$ with $\alpha > 0$. Regarding the out-degrees, we observed that the lower out-degrees deviate significantly from a power law distribution but the larger out-degrees again approximately fit such a distribution. A similar deviation of the low out-degrees was observed by Broder et al. [Broder et al. 00] for a sample of the web.

In our experiment on the bibliographic data, we trace the authority of authors over the last five decades, i.e., the 1960s, 1970s, 1980s, 1990s, and 2000s. For each decade we define the temporal interest, so that the temporal window of interest is on the decade itself and the tolerance interval includes two years before and after the decade. The temporal interest employed for the 1980s, as an example, uses the following parameter setting: $t_1 = \text{January 1st, 1978}$, $TS_{Origin} =$

January 1st, 1980, $TS_{End} =$ December 31st, 1989, and $t_2 =$ December 31st, 1991. Because the current decade is not yet finished, we employ $TS_{End} = t_2 =$ December 31st, 2004, for the upper limits of the temporal window of interest and tolerance interval. The small constant e that was previously used for the definition of freshness and activity is assumed as $e = 10^{-10}$. The parameters of PageRank, T-Rank Light, and T-Rank are set as follows: For the random jump probability, we assume $\epsilon = 0.15$. The threshold δ giving the stopping criterion for the power iteration method is assumed as $\delta = 10^{-10}$. For the coefficients of the time-aware methods, we employ the uniform parameter setting $w_{si} = 1/4$ and $w_{ti} = 1/6$. Before presenting results obtained on the bibliographic data, we clearly state that these results must be interpreted with caution. Thus, the presented lists do not necessarily capture the true scientific merits of the authors but merely reflect the authority that is expressed by the fairly partial citation information available in DBLP.

To provide the reader with an insight into our results, we present in Table 1 the top-15 lists obtained from PageRank, T-Rank Light, and T-Rank for the 2000s. Given this temporal interest, we expect the time-aware methods to bring to the top positions authors who have many publications and are intensively cited in the 2000s, thus yielding high freshness and activity values for the corresponding nodes and high average freshness and activity for the edges pointing to these nodes.

Comparing the depicted top-15 lists, we observe that the overlap between the top-15 lists produced by PageRank and T-Rank Light is larger than the overlap between the top-15 lists obtained from PageRank and T-Rank. This is due to the fact that T-Rank's random walk is more biased than that of T-Rank Light. From the depicted top-15 lists, one can draw the following anecdotic evidence for the time-aware methods producing indeed meaningful rankings. PageRank brings E. F. Codd to the top position, although this author has no publications recorded in DBLP after the year 1990, leading to low freshness and activity of the corresponding node in the graph. Thus, it is intuitive that the author appears only at the second position in the top-15 list obtained from T-Rank Light and not at all in the top-15 list produced by T-Rank. Considering our prior remark that the time-aware methods should bring up authors who have many highly cited publications in the 2000s, it is not astonishing that, for example, Hector Garcia-Molina, Jennifer Widom, and Rakesh Agrawal appear in the time-aware top-15 lists, since all these authors have publications in the 2000s that are highly cited according to DBLP.

Moreover, we compared the rankings obtained from the ranking methods for consecutive decades. Thus, we investigate the time-sensitivity of the methods, i.e., how intensively the methods react to the different temporal interests and the

	PageRank	T-Rank Light	T-Rank
1	E. F. Codd	Michael Stonebraker	Jim Gray
2	Michael Stonebraker	E. F. Codd	Michael Stonebraker
3	Jim Gray	Jim Gray	Jeffrey D. Ullman
4	Donald D. Chamberlin	Jeffrey D. Ullman	Philip A. Bernstein
5	Jeffrey D. Ullman	Donald D. Chamberlin	Hector Garcia-Molina
6	Philip A. Bernstein	Philip A. Bernstein	Jeffrey F. Naughton
7	Raymond A. Lorie	Raymond A. Lorie	Donald D. Chamberlin
8	Morton M. Astrahan	Morton M. Astrahan	David J. DeWitt
9	Kapali P. Eswaran	Kapali P. Eswaran	Jennifer Widom
10	John Miles Smith	Irving L. Traiger	Rakesh Agrawal
11	Irving L. Traiger	David J. DeWitt	Hamid Pirahesh
12	Peter P. Chen	John Miles Smith	Joseph M. Hellerstein
13	Eugene Wong	Nathan Goodman	Yehoshua Sagiv
14	Nathan Goodman	Eugene Wong	Stefano Ceri
15	David J. DeWitt	Hector Garcia-Molina	Andrew Eisenberg

Table 1. Top-15 lists produced by PageRank, T-Rank Light, and T-Rank for the 2000s.

changes in the underlying evolving graph. For this investigation, we employed the following similarity measures for the comparison of top- k lists. Let L_1 and L_2 be two top- k lists over a set S of n objects ($k < n$):

1. the overlap similarity **OSim** [Haveliwala 02], which is the fraction of objects that appears in both top- k lists, i.e., $OSim(L_1, L_2) = |L_1 \cap L_2|/k$.
2. the **KSim** similarity [Haveliwala 02], which is based on Kendall's τ and describes the probability that the two top- k lists agree on the relative order of a randomly selected pair $(u, v) \in S \times S$.

All comparisons in the context of this experiment on bibliographic data were conducted on the basis of top-1,000 lists. Thus, when referring to a comparison of rankings, we implicitly refer to the comparison of the corresponding top-1,000 lists.

The similarities obtained for the comparison of the consecutive decades are depicted in Figure 3. With the exception of the KSim similarities obtained for the comparison of the 1970s and 1980s, we observe that the time-aware methods T-Rank Light and T-Rank accomplish consistently lower similarities for consecutive decades than PageRank. Thus, on the basis of the modified temporal interest and the changes in the underlying graph structure, the time-aware methods react more intensively to the modified temporal interest and the changes in the underlying graph structure. Interestingly, we observe that for the comparison of the 1990s and 2000s, PageRank does not produce noticeably different rankings, which is due to the fact that there are few newly recorded citations after the

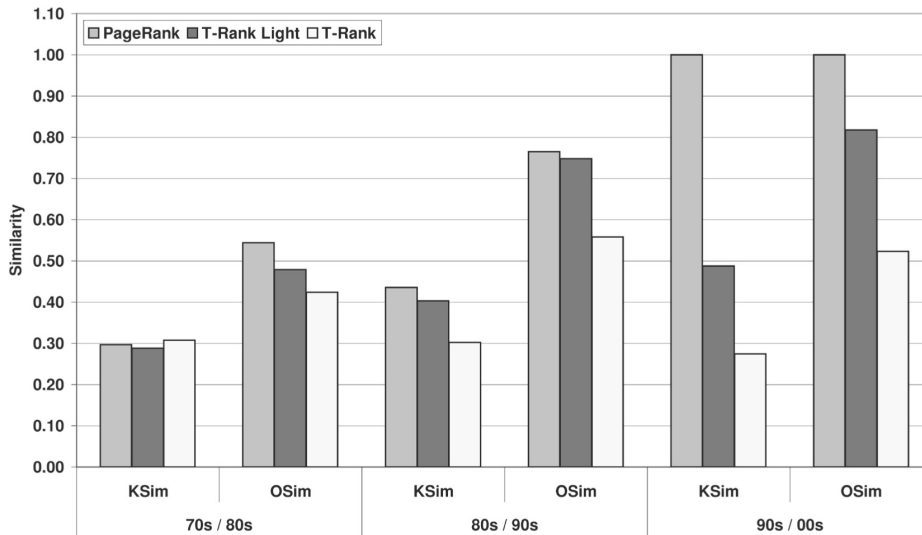


Figure 3. Similarities between rankings obtained from the same ranking method for consecutive decades.

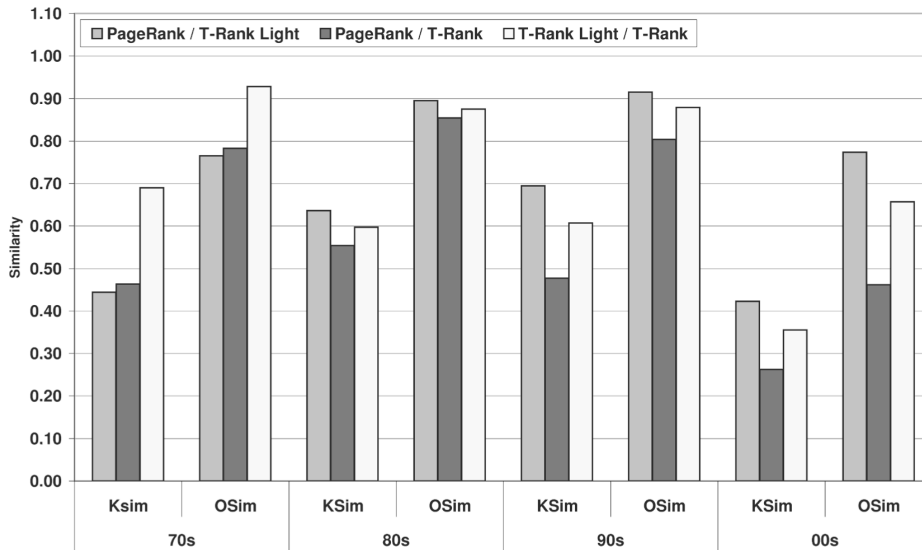


Figure 4. Similarities between rankings obtained from different ranking methods for the same decade.

year 2000 and thus only little changes to the connected part of the underlying graph structure. Nevertheless, the time-aware methods still produce significantly different rankings for the two decades, since they take into account also the newly recorded publications (i.e., modifications to the nodes in the graph) that have no effect on PageRank. Another finding drawn from the depicted similarities is that the similarities observed for T-Rank Light and T-Rank vary less than those observed for PageRank. Regarding the time-aware methods, similarities for T-Rank differ less than those observed for T-Rank Light. This finding indicates that the time-aware methods are more robust than PageRank.

Furthermore, we investigated the similarities between rankings obtained from the different ranking methods for the same decade. The resulting similarity scores are given in Figure 4. From the depicted figures we observe that, with the exception of the 1970s, the rankings obtained from PageRank exhibits higher similarity to the ranking produced by T-Rank Light than to the ranking produced by T-Rank. We consider this finding intuitive, regarding that also in terms of the underlying random walk, T-Rank Light is closer to PageRank, since it biases only the random jump. Comparing the time-aware rankings, we observe that the rankings produced by T-Rank are closer to the rankings produced by T-Rank Light than to the rankings obtained from PageRank, which is also intuitive, regarding that the two time-aware methods were defined in an incremental manner.

4.2. Experiment on Amazon.com Product Data

The second dataset contains information about approximately 200,000 products (mostly books) offered by Amazon.com. The data was obtained through a web service provided by Amazon.com [Amazon 05], but it can be regarded as being equivalent to a partial crawl of the corresponding web site. We chose this way of obtaining the data for the sake of easier processing and better data quality. We map the products to nodes in our evolving graph structure and obtain the timestamp $TS_{Creation}$ as the publication date of the respective product. Timestamps in $TS_{Modifications}$ are derived from the dates when a customer reviewed the product. For each product, pointers to similar recommended products are available in the dataset. These pointers define the edges in our graph. For their temporal annotations we take the maximal timestamp $TS_{Creation}$ of the source and the target node, i.e., the publication dates of the similar products, so that the edge is supposed to be created with the younger of the two nodes. The dataset was obtained between February 2005 and April 2005, so the majority of temporal annotations date to the time before April 2005. However, since some

products are already listed before their actual publication, there are temporal annotations dating to times in the future.

We applied PageRank, T-Rank Light, and T-Rank on this dataset. The probability of making a random jump was set to $\epsilon = 0.15$, and the convergence criterion was assumed to be $\delta = 10^{-10}$. For the temporal interest, we chose a tolerance interval including the last quarter in 2004 as well as the first quarter in 2005, i.e., we set $t_1 = \text{October 1st, 2004}$, and $t_2 = \text{March 31st, 2005}$. The temporal window of interest was set to focus on the first quarter in 2005, i.e., we chose $TS_{Origin} = \text{January 1st, 2005}$, and $TS_{End} = \text{March 31st, 2005}$. For the parameters w_{si} and w_{ti} of the time-aware methods, uniform parameter assignments were employed, i.e., these parameters were assumed as $w_{si} = 1/4$ and $w_{ti} = 1/6$.

On this basis we computed rankings employing PageRank, T-Rank Light, and T-Rank. Using these precomputed rankings, we produced top-ten lists for 30 different queries. For each of the queries, three different top-ten lists were produced ordering the contained products only according to the respective ranking (i.e., no content-based ranking such as tf-idf variants or Okapi BM25 was used). Products whose associated textual data contained all of the query terms were regarded as relevant to a specific query. For one third of the queries, matches were identified solely based on the associated customer reviews, the associated editorial reviews, and the product titles, respectively. In all cases, stemming of both the textual data and the query was used to increase the number of qualifying documents.

We conducted a user study to evaluate the quality of the obtained top-ten lists. For this user study, five test users were asked to grade the quality of the three different top-ten lists for each of the thirty queries. To this end, for every query the three different top-ten lists were labeled as “Method 1,” “Method 2,” and “Method 3” thus making the employed ranking method anonymous. The test users could employ grades from 0 (worst grade) to 3 (best grade) for the assessment. For every query and every ranking method, we summed up and normalized the grades that were given by our test users. The aggregate grades obtained on the individual queries are depicted in Figure 5. In the figure the queries are marked with the letters “C,” “E,” and “T” to indicate whether customer reviews, editorial reviews, or product titles were used for identifying the query matches.

From Figure 5 it can be seen that PageRank produced the best graded top-ten list for seven of the queries. For three queries (i.e., “presidential election 2004,” “summer gardening,” and “easy cooking”), the gradings of the top-ten lists produced by PageRank lie between the better and the worse of the top-ten lists produced by the time-aware methods. Finally, for 20 of the queries both

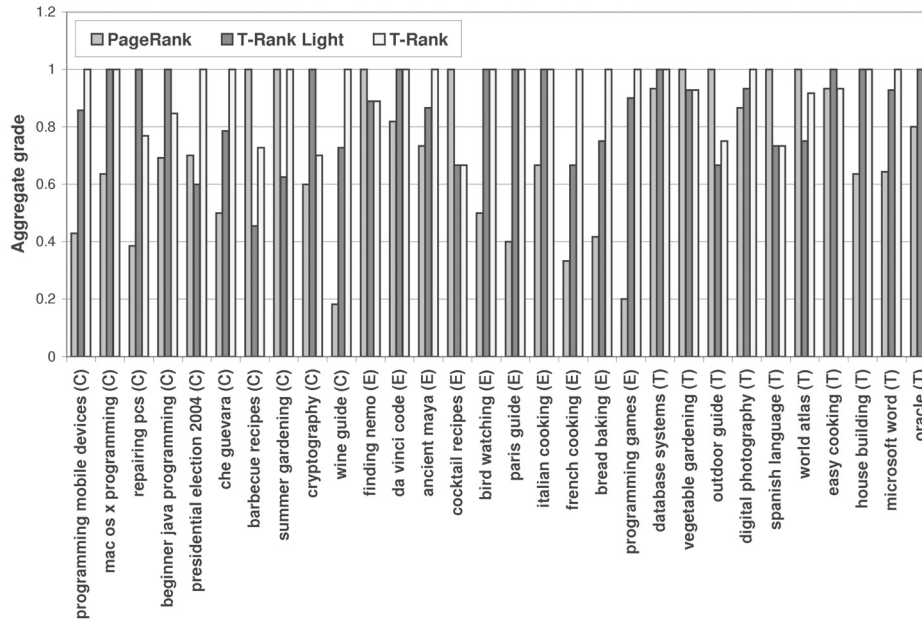


Figure 5. Aggregate grades of PageRank, T-Rank Light, and T-Rank for individual queries.

T-Rank Light and T-Rank produced top-ten lists that received better grades from our test users than the top-ten obtained from PageRank. So, for the clear majority of the queries, i.e., more than 66%, the top-ten lists produced by the time-aware methods are regarded as superior to the top-ten list produced by PageRank.

Comparing the grades given to T-Rank Light and T-Rank, we observe that for 14 queries, the top-ten list produced by T-Rank is graded better than the top-ten list produced by T-Rank Light. Ties between the two time-aware methods are observed for twelve queries and, finally, the top-ten list produced by T-Rank Light is graded better than the top-ten list obtained from T-Rank for the four queries “repairing pcs,” “beginner java programming,” “cryptography,” and “easy cooking.” Thus, as a conclusion, for nearly half of the queries, the additional effort, i.e., the skewing of the transition probabilities, inherent to T-Rank results in a better top-ten list, whereas a deterioration of the top-ten list is only observed for four of the queries.

For illustration of the results, Table 2 shows the top-five lists produced by PageRank, T-Rank Light, and T-Rank for the query “programming games.” For the top-five list produced by PageRank, we regard only one of the products

Top-five returned by PageRank	
1	Macromedia Flash MX 2004 ActionScript: Training from the Source http://www.amazon.com/exec/obidos/tg/detail/-/0321213432
2	Microeconomic Analysis http://www.amazon.com/exec/obidos/tg/detail/-/0393957357
3	Game Programming Gems 4 (Game Programming Gems Series) http://www.amazon.com/exec/obidos/tg/detail/-/1584502959
4	Five-Star Basketball Drills http://www.amazon.com/exec/obidos/tg/detail/-/0940279223
5	Complete Conditioning for Tennis http://www.amazon.com/exec/obidos/tg/detail/-/0880117346
Top-five returned by T-Rank Light	
1	Patterns in Game Design (Game Development Series) http://www.amazon.com/exec/obidos/tg/detail/-/1584503548
2	Game Level Design (Game Development Series) http://www.amazon.com/exec/obidos/tg/detail/-/1584503696
3	Real-Time Cinematography for Games (Game Development Series) http://www.amazon.com/exec/obidos/tg/detail/-/1584503084
4	Macromedia Flash MX 2004 ActionScript: Training from the Source http://www.amazon.com/exec/obidos/tg/detail/-/0321213432
5	Game Programming Gems 4 (Game Programming Gems Series) http://www.amazon.com/exec/obidos/tg/detail/-/1584502959
Top-five returned by T-Rank	
1	Patterns in Game Design (Game Development Series) http://www.amazon.com/exec/obidos/tg/detail/-/1584503548
2	Game Level Design (Game Development Series) http://www.amazon.com/exec/obidos/tg/detail/-/1584503696
3	Beginning 3D Game Programming http://www.amazon.com/exec/obidos/tg/detail/-/0672326612
4	Programming a Multiplayer FPS in DirectX http://www.amazon.com/exec/obidos/tg/detail/-/1584503637
5	Real-Time Cinematography for Games (Game Development Series) http://www.amazon.com/exec/obidos/tg/detail/-/1584503084

Table 2. Top-five lists produced by PageRank, T-Rank Light, and T-Rank for the query “programming games.”

(“Game Programming Gems 4”) as a good result. The top-five lists obtained from the time-aware methods T-Rank Light and T-Rank, on the other hand, contain four and five products, respectively, that are indeed good matches for the given query.

4.3. Experiment on Web Data

We conducted a third series of experiments on web data and investigated in a user study how the time-aware rankings are accepted by the users. As a dataset for this experiment, we chose approximately 200,000 pages related to the 2004 Olympic Games and crawled them repeatedly before, during, and after the Olympic Games. More precisely, we fixed the set of documents by conducting an initial crawl (on July 26, 2004) that had the top-100 lists returned by Google.com for the queries “Olympic Games,” “Athens 2004,” and “Athens Olympics” as its starting points. In total, we conducted nine observations of the documents in this set and on this basis derived an evolving graph.

Based on this evolving graph, we applied PageRank, T-Rank Light, and T-Rank employing the following parameter setting: The tolerance interval of the temporal interest was set to cover the period of our crawls, i.e., we set $t_1 = \text{July 26, 2004}$ and $t_2 = \text{September 1, 2004}$; the temporal window of interest was set to include only the period of the Olympic Games, i.e., $TS_{Origin} = \text{August 14, 2004}$ and $TS_{End} = \text{August 31, 2004}$; the random jump probability was assumed as $\epsilon = 0.15$ and the stopping threshold of the power iteration method was assumed as $\delta = 10^{-10}$.

For the user study, we selected six queries and computed top-ten lists according to PageRank, T-Rank Light, and T-Rank. The top-ten lists were computed as lists of documents containing all keywords and having the ten highest authority scores. The user study was conducted in September, so we could profit from Google.com’s Zeitgeist service [Google 05]. This service presents queries that were frequently asked within the past month or that are related to some event. Half of our queries are drawn from this service and are therefore known to have been frequently asked during the Olympic Games.

In a first assessment, we asked users to grade the overall quality of top-ten lists presented for a specific query. To this end, the top-ten lists were labeled as “Method 1,” “Method 2,” and “Method 3” thus making the employed ranking method anonymous. For the assessment, users could give grades from 1 (worst grade) to 3 (best grade). The second assessment presented to the users a list of at most 30 documents (i.e., the shuffled distinct union of the top-ten lists produced by the ranking methods) and asked them to grade the relevance of the single documents with regard to a given query. Here, users could assign grades from 0 (worst score) to 5 (best score) for the assessment. Based on the users’ assessments, for every query, we computed a users’ top-ten list (i.e., the ten documents with highest aggregate grades), which reflects the users’ perceptions for the specific query. Employing the similarity measures introduced in Section 4.1, we compared this users’ top-ten list to the top-ten lists obtained from the ranking methods.

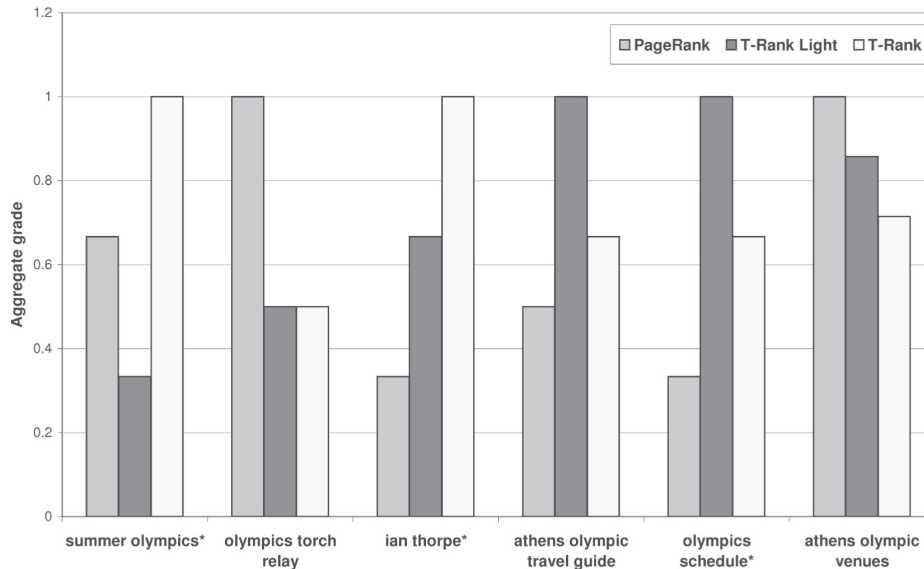


Figure 6. Aggregate grades of PageRank, T-Rank Light, and T-Rank for individual queries.

Figure 6 presents the results, i.e., the aggregate grades on a per query basis, of the first assessment. Those queries drawn from Google.com Zeitgeist are indicated with an asterisk. As depicted in Figure 6, PageRank produces the top-ten list favored by the users for the two queries “olympics torch relay” and “athens olympic venues.” For the other four queries one of the time-aware methods produces the preferred top-ten list. Altogether, for half of the queries both of the time-aware methods deliver a top-ten list that is superior to the one obtained from PageRank, and for another half of the queries, PageRank outperforms either T-Rank Light or T-Rank. Thus, the coarse-grained evaluation of the ranking methods conducted in this first assessment does not clearly indicate any superiority of the time-aware methods.

The results of the second assessment for which the users were asked in our study is illustrated in Figure 7. The figure depicts the OSim and KSim similarities between the users’ top-ten lists and the top-ten lists produced by PageRank, T-Rank Light, and T-Rank. In this second assessment PageRank is clearly ahead of the time-aware methods only for the query “olympics torch relay” and achieves the same highest overlap as T-Rank Light for the query “athens olympic travel guide.” For the other four queries, both time-aware methods produce top-ten lists that have higher similarity to the users’ top-ten list than the top-ten list

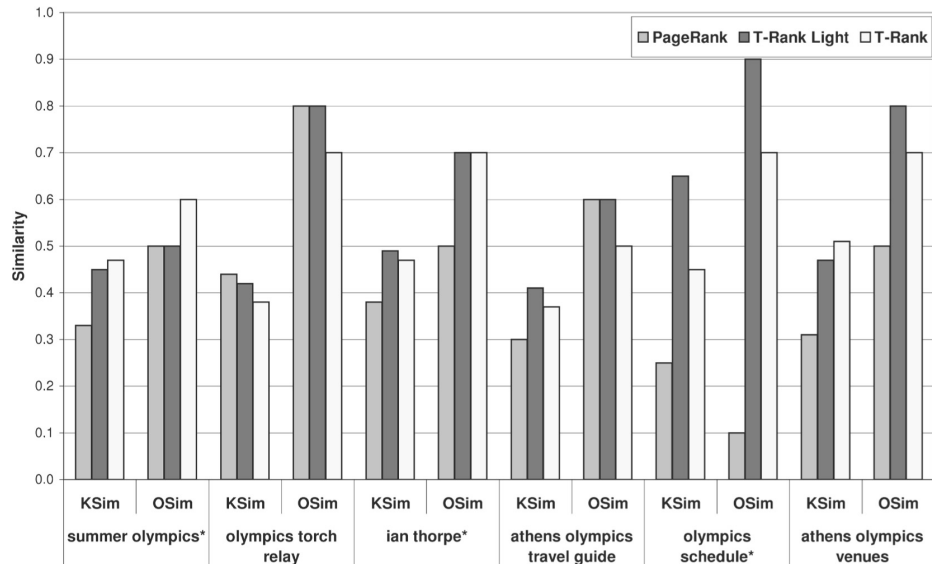


Figure 7. Similarities between users' top-ten lists and top-ten lists produced by ranking methods.

delivered by PageRank. Thus, in this second, more fine-grained assessment, for the majority of the considered queries, both time-aware methods produce top-ten lists that are equally good or better than the top-ten list produced by PageRank. Comparing the time-aware methods, we observe that, again, we cannot claim that either of the time-aware methods outperforms the other.

To provide the reader with anecdotic insight into the results produced by PageRank and the time-aware methods, we pick out the query “ian thorpe,” which is one of the queries that was frequently submitted by users during the Olympic Games, and present the users' top-ten list as well as the top-ten obtained from the ranking methods in Table 3 and Table 4.

From the presented results it is visible that PageRank accomplishes an overlap of five documents with the the users' top-ten lists, whereas both time-aware methods achieve overlaps of seven documents agreeing with the OSIm values given in Figure 7. In addition, we make the following observation: among the top-ten list produced by PageRank, there are three documents (entitled “ISOH—International Society of Olympic Historians,” “2000 Olympics,” and “World Olympic Association”) that—as we verified—are not directly related to the 2004 Olympic Games. For the time-aware methods, only one of these documents, namely the page entitled “2000 Olympics,” shows up. Furthermore, among the users' top-ten list there are four news articles that appeared in the Sydney Morn-

Top-ten aggregated from user assessment	
1	Is there anything Thorpe can't do? - Swimming http://www.smh.com.au/olympics/articles/2004/07/23/1090464859991.html
2	BBC SPORT — Commonwealth Games 2002 — Front Page http://news.bbc.co.uk/sport3/commonwealthgames2002/default.stm
3	NBC Olympics - Athlete Bios http://nbcolympics.com/athletebios
4	Guardian Unlimited Sport — Special reports http://sport.guardian.co.uk/athens2004
5	The must-see moments of the Athens Games - Aussie Update http://www.smh.com.au/olympics/articles/2004/07/23/1090464860281.html
6	Dreams can differ from Olympic reality http://www.smh.com.au/olympics/articles/2004/07/23/1090464859962.html
7	Pressure to beat Americans could sink swimmers http://www.smh.com.au/olympics/articles/2004/07/25/1090693838106.html
8	MSNBC Newsweek Olympics Front Page http://www.msnbc.msn.com/id/5112034/site/newsweek
9	NBC Olympics - Swimming http://www.nbcolympics.com/swimming
10	NBC Olympics - Athlete Bios Michael Phelps http://www.nbcolympics.com/athletebios/5009163/detail.html
Top-ten returned by PageRank	
1	NBC Olympics - Swimming http://www.nbcolympics.com/swimming
2	NBC Olympics - Athlete Bios http://nbcolympics.com/athletebios
3	ISOH - International Society of Olympic Historians http://www.olykamp.org/isoh/membership-alpha.html
4	North County Times - North San Diego and Southwest Riverside County http://www.nctimes.com/special_reports/olympics
5	2000 Olympics http://print.infoplease.com/ipsa/A0922268.html
6	BBC SPORT — Olympics 2004 — Britain's full Olympic line-up http://news.bbc.co.uk/sport1/hi/olympics_2004/3685595.stm
7	BBC SPORT — Commonwealth Games 2002 — Front Page http://news.bbc.co.uk/sport3/commonwealthgames2002/default.stm
8	NBC Olympics - Athlete Bios Michael Phelps http://www.nbcolympics.com/athletebios/5009163/detail.html
9	Guardian Unlimited Sport — Special reports http://sport.guardian.co.uk/athens2004
10	World Olympics Association http://www.woaolympians.com/agree_aid.html

Table 3. Users' top-ten list and top-ten list produced by PageRank.

ing Herald (<http://www.smh.com.au>) covering the 2004 Olympic Games. None of these articles appear in the top-ten list obtained from PageRank. The time-aware methods T-Rank Light and T-Rank, on the other hand, yield three and four of these relevant recent documents, respectively. Thus, as an overall conclusion of this anecdotic example, the time-aware methods better capture the users' temporal interest on the 2004 Olympic Games.

Top-ten returned by T-Rank Light	
1	NBC Olympics - Swimming http://www.nbcolympics.com/swimming
2	NBC Olympics - Athlete Bios http://nbcolympics.com/athletebios
3	North County Times - North San Diego and Southwest Riverside County http://www.nctimes.com/special_reports/olympics
4	MSNBC Newsweek Olympics Front Page http://www.msnbc.msn.com/id/5112034/site/newsweek
5	BBC SPORT — Olympics 2004 — Britain's full Olympic line-up http://news.bbc.co.uk/sport1/hi/olympics_2004/3685595.stm
6	NBC Olympics - Athlete Bios Michael Phelps http://www.nbcolympics.com/athletebios/5009163/detail.html
7	Is there anything Thorpe can't do? - Swimming http://www.smh.com.au/olympics/articles/2004/07/23/1090464859991.html
8	2000 Olympics http://www.infoplease.com/ipsa/A0922268.html
9	Pressure to beat Americans could sink swimmers http://www.smh.com.au/olympics/articles/2004/07/25/1090693838106.html
10	Dreams can differ from Olympic reality http://www.smh.com.au/olympics/articles/2004/07/23/1090464859962.html
Top-ten returned by T-Rank	
1	NBC Olympics - Athlete Bios http://nbcolympics.com/athletebios
2	Olympic NBC Olympics - Swimming http://www.nbcolympics.com/swimming
3	North County Times - North San Diego and Southwest Riverside County http://www.nctimes.com/special_reports/olympics
4	BBC SPORT — Olympics 2004 — Britain's full Olympic line-up http://news.bbc.co.uk/sport1/hi/olympics_2004/3685595.stm
5	NBC Olympics - Athlete Bios Michael Phelps http://www.nbcolympics.com/athletebios/5009163/detail.html
6	Is there anything Thorpe can't do? - Swimming http://www.smh.com.au/olympics/articles/2004/07/23/1090464859991.html
7	Pressure to beat Americans could sink swimmers http://www.smh.com.au/olympics/articles/2004/07/25/1090693838106.html
8	2000 Olympics http://www.infoplease.com/ipsa/A0922268.html
9	Dreams can differ from Olympic reality http://www.smh.com.au/olympics/articles/2004/07/23/1090464859962.html
10	The must-see moments of the Athens Games - Aussie Update http://www.smh.com.au/olympics/articles/2004/07/23/1090464860281.html

Table 4. Top-ten lists produced by T-Rank Light and by T-Rank.

4.4. Parameter Sensitivity Study

We investigated the parameter sensitivity of the time-aware methods with regard to the parameters w_{si} and w_{ti} on all considered datasets. For the bibliographic dataset we chose the 1990s as a representative decade that has plenty of recorded publications and citations. For both the Amazon.com and the web dataset, the

same temporal interests as in the previous experiments were used. The remaining parameters were set as follows: $\epsilon = 0.15$ and $\delta = 10^{-10}$. All comparisons of rankings in this section were conducted on the basis of top-1,000 lists. Thus, again, when referring to a comparison of rankings, we refer to the comparison of the corresponding top-1,000 lists.

First, we analyzed the parameter sensitivity of T-Rank Light. We computed a baseline ranking (TRL_B) employing the uniform parameter setting $w_{si} = 1/4$. Furthermore, we computed additional rankings that emphasize freshness (TRL_F), activity (TRL_A), temporal information on nodes (TRL_N), and average temporal information on incoming edges (TRL_{IE}). The detailed parameter settings for those rankings are given in Table 5. We compared the rankings produced by nonuniform parameter settings to the baseline rankings. The similarities obtained on the different datasets are given in Table 6. From the figures we observe the following: T-Rank Light produces rankings that are more similar to the baseline ranking if all weight is put on freshness or activity. With the exception of the ranking TRL_{IE} produced on the bibliographic dataset, rankings with lower similarity to the baseline ranking are obtained if temporal information on nodes or average temporal information on incoming edges is emphasized.

Moreover, we compared the two rankings emphasizing freshness and activity, as well as the two rankings emphasizing temporal information on nodes and average temporal information on incoming edges. Table 7 lists the obtained similarities. The figures consistently indicate the following: T-Rank Light produces more similar rankings if emphasis is shifted between freshness and activity, thus

	TRL_B	TRL_F	TRL_A	TRL_N	TRL_{IE}
w_{s1}	1/4	1/2	0	1/2	0
w_{s2}	1/4	1/2	0	0	1/2
w_{s3}	1/4	0	1/2	1/2	0
w_{s4}	1/4	0	1/2	0	1/2

Table 5. Parameter settings employed for analysis of T-Rank Light.

		TRL_B/TRL_F	TRL_B/TRL_A	TRL_B/TRL_N	TRL_B/TRL_{IE}
DBLP	KSim	0.87	0.84	0.54	0.92
	OSim	0.97	0.96	0.85	0.98
Amazon	KSim	0.62	0.59	0.28	0.46
	OSim	0.89	0.89	0.54	0.78
Web	KSim	0.82	0.79	0.56	0.54
	OSim	0.96	0.95	0.87	0.85

Table 6. Similarities between ranking produced by uniform parameter settings and rankings produced by nonuniform parameter settings (T-Rank Light).

		TRL_F/TRL_A	TRL_N/TRL_{IE}
DBLP	KSim	0.74	0.51
	OSim	0.94	0.83
Amazon	KSim	0.44	0.28
	OSim	0.78	0.33
Web	KSim	0.67	0.38
	OSim	0.91	0.74

Table 7. Similarities between rankings obtained from nonuniform parameter setting (T-Rank Light).

	TR_B	TR_F	TR_A	TR_N	TR_{SE}	TR_{IE}
w_{t1}	1/6	1/3	0	1/2	0	0
w_{t2}	1/6	1/3	0	0	1/2	0
w_{t3}	1/6	1/3	0	0	0	1/2
w_{t4}	1/6	0	1/3	1/2	0	0
w_{t5}	1/6	0	1/3	0	1/2	0
w_{t6}	1/6	0	1/3	0	0	1/2

Table 8. Parameter settings employed for analysis of T-Rank.

reacting less intensively. On the contrary, less similar rankings are obtained if emphasis is shifted between the temporal information on nodes and the average temporal information on incoming edges, which indicates a stronger sensitivity of T-Rank Light.

Second, we investigated the parameter sensitivity of T-Rank. To this end, we fixed the parameters affecting the random jump to the uniform parameter setting $w_{si} = 1/4$. The rationale here is that we expect our findings obtained for T-Rank Light to transfer to T-Rank. Then, we varied the setting of the parameters w_{ti} that affect the transition probabilities. Employing the uniform parameter setting $w_{ti} = 1/6$, we computed a baseline ranking (TR_B). Additionally, we computed five rankings that put emphasis on freshness (TR_F), activity (TR_A), temporal information on nodes (TR_N), temporal information on individual edges (TR_{SE}), and average temporal information on incoming edges (TR_{IE}). Table 8 presents the detailed parameter settings employed for these rankings. For the parameter sensitivity study, in a first step, we compared the rankings produced by nonuniform parameter settings to the baseline ranking. The resulting similarities are given in Table 9. On this basis we observe that, on our datasets, the rankings emphasizing only freshness or activity are most similar to the baseline ranking. Moreover, the ranking with the lowest similarity to the baseline ranking is the ranking emphasizing the temporal information on nodes.

		TR_B/TR_F	TR_B/TR_A	TR_B/TR_N	TR_B/TR_{SE}	TR_B/TR_{IE}
DBLP	KSim	0.78	0.77	0.61	0.69	0.74
	OSim	0.95	0.95	0.88	0.92	0.94
Amazon	KSim	0.96	0.97	0.39	0.49	0.62
	OSim	0.99	0.99	0.71	0.82	0.89
Web	KSim	0.80	0.80	0.41	0.48	0.49
	OSim	0.96	0.96	0.71	0.83	0.83

Table 9. Similarities between baseline ranking and rankings obtained from skewed parameter settings (T-Rank).

		TR_F/TR_A	TR_N/TR_{SE}	TR_N/TR_{IE}	TR_{SE}/TR_{IE}
DBLP	KSim	0.63	0.49	0.52	0.77
	OSim	0.91	0.81	0.83	0.94
Amazon	KSim	0.94	0.34	0.46	0.40
	OSim	0.98	0.59	0.79	0.72
Web	KSim	0.66	0.33	0.33	0.52
	OSim	0.91	0.57	0.58	0.84

Table 10. Similarities between rankings obtained from skewed parameter settings (T-Rank).

For the rankings putting all weight on temporal information on edges, slightly higher similarity to the baseline ranking is observed for the ranking TR_{IE} , i.e., the ranking emphasizing average temporal information on incoming edges.

We also compared the two rankings emphasizing freshness and activity, as well as all pairs of rankings emphasizing temporal information on nodes or edges. The resulting similarities are presented in Table 10. Here, we notice high similarity for the rankings emphasizing freshness or activity and high similarity for the rankings putting weight on any kind of temporal information on edges. In contrast, the rankings emphasizing the temporal information on nodes have comparably low similarity to the rankings putting weight on temporal information on edges.

Thus, as a conclusion from the parameter sensitivity study, we claim the following. On the test datasets, both T-Rank Light and T-Rank are sensitive to changes of weights for temporal information on nodes, temporal information on single edges, or aggregate temporal information on incoming edges. On the other hand, we observe much lower sensitivity when the influence of freshness and activity is modified. So, with regard to the latter parameters, T-Rank Light and T-Rank are fairly robust methods.

5. Conclusions

In this paper we devised two time-aware authority ranking methods, which can both be seen as variants of the PageRank method. T-Rank Light skews only the random jump probabilities, whereas T-Rank is more comprehensive by additionally skewing the transition probabilities. Given a user's temporal interest, both methods consider the temporal aspects freshness and activity of pages and links to produce time-aware authority rankings. In our experiments on bibliographic data, the methods produced meaningful rankings. In our second experiment on the data obtained from Amazon.com, a user study revealed both time-aware methods to produce top-ten lists superior to the top-ten lists produced by PageRank. The user study conducted as a part of our third experiment, which was run on a subset of the web, eventually gave further evidence that the time-aware methods produce rankings that are indeed favored by the users. An analysis of the parameter sensitivity revealed that both time-aware methods are robust with regard to the choice of freshness-vs.-activity parameters, whereas they are more sensitive to the parameters that control the influence of nodes, edges, and aggregated incoming edges.

Currently, there is a movement towards archiving web contents, with the Internet Archive [Archive 05] being the most prominent project in this context. The size of a comprehensive web archive would be vast, but such an archive would be a valuable corpus for queries with strong temporal flavor. For the experiments on web data presented in this paper, we approximated such an archive by repeatedly crawling a small subset of the web. Regardless of how comprehensive the available evolutionary information is, the user's temporal interest should be taken into account for the ranking of web pages. Thus, we believe that time-awareness is a very important issue in the quest for meaningful ranking techniques in web search.

References

- [Amazon 05] Amazon.com. "Web Services." Available from World Wide Web (<http://www.amazon.com/gp/aws/landing.html>), 2005.
- [Amitay et al. 04] Einat Amitay, David Carmel, Michael Herscovici, Ronny Lempel, and Aya Soffer. "Trend Detection Through Temporal Link Analysis." *Journal of the American Society for Information Science & Technology* 55:14 (2004), 1270–1281.
- [Archive 05] Internet Archive home page. Available from World Wide Web (<http://www.archive.org>), 2005.

- [Baeza-Yates et al. 02] Ricardo A. Baeza-Yates, Felipe Saint-Jean, and Carlos Castillo. “Web Structure, Dynamics and Page Quality.” In *Proceedings of the 9th International Symposium on String Processing and Information Retrieval*, pp. 117–130. London: Springer-Verlag, 2002.
- [Baeza-Yates et al. 04] Ricardo A. Baeza-Yates, Carlos Castillo, and Felipe Saint-Jean. “Web Structure, Dynamics and Page Quality.” In *Web Dynamics*, edited by Mark Levene and Alexandra Poulouvassilis, pp. 93–115. New York: Springer-Verlag, 2004.
- [Bar-Yossef et al. 04] Ziv Bar-Yossef, Andrei Z. Broder, Ravi Kumar, and Andrew Tomkins. “Sic Transit Gloria Telae: Towards an Understanding of the Web’s Decay.” In *Proceedings of the 13th International Conference on World Wide Web*, pp. 328–337. New York: ACM Press, 2004.
- [Brin and Page 98] Sergey Brin and Lawrence Page. “The Anatomy of a Large-Scale Hypertextual Web Search Engine.” *Computer Networks and ISDN Systems* 30:1–7 (1998), 107–117.
- [Broder et al. 00] Andrei Broder, Ravi Kumar, Farzin Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins, and Janet Wiener. “Graph Structure in the Web.” In *Proceedings of the 9th International World Wide Web Conference, Computer Networks* 33:1–6 (2000), 309–320.
- [Fetterly et al. 03] Dennis Fetterly, Mark Manasse, Marc Najork, and Janet Wiener. “A Large-Scale Study of the Evolution of Web Pages.” In *Proceedings of the 12th International Conference on World Wide Web*, pp. 669–678. New York: ACM Press, 2003.
- [Google 05] Google, Inc. “Google Zeitgeist: Search Patterns, Trends, and Surprises According to Google.” Available from World Wide Web (<http://www.google.com/press/zeitgeist.html>), 2005.
- [Haveliwala 02] Taher H. Haveliwala. “Topic-Sensitive PageRank.” In *Proceedings of the 11th International Conference on World Wide Web*, pp. 517–526. New York: ACM Press, 2002.
- [Jeh and Widom 03] Glen Jeh and Jennifer Widom. “Scaling Personalized Web Search.” In *Proceedings of the 12th International Conference on World Wide Web*, pp. 271–279. New York: ACM Press, 2003.
- [Kleinberg 99] Jon M. Kleinberg. “Authoritative Sources in a Hyperlinked Environment.” *Journal of the ACM* 46:5 (1999), 604–632.
- [Kraft et al. 03] Reiner Kraft, Enes Hastor, and Raymie Stata. “TimeLinks: Exploring the Link Structure of the Evolving Web.” In *WAW2003 Second Workshop on Algorithms and Models for the Web-Graph*. Available from World Wide Web (<http://www.soe.ucsc.edu/~rekrft/papers/workshop.pdf>), 2003.
- [Kumar et al. 00] Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, D. Sivakumar, Andrew Tompkins, and Eli Upfal. “The Web as a Graph.” In *Proceedings of the Nineteenth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, pp. 1–10. New York: ACM Press, 2000.
- [Kumar et al. 05] Ravi Kumar, Jasmine Novak, Prabhakar Raghavan, and Andrew Tomkins. “On the Bursty Evolution of Blogspace.” *World Wide Web* 8:2 (2005), 159–178.

- [Lempel and Moran 01] Ronny Lempel and Shlomo Moran. “SALSA: The Stochastic Approach for Link-Structure Analysis.” *ACM Transactions on Information Systems* 19:2 (2001), 131–160.
- [Ntoulas et al. 04] Alexandros Ntoulas, Junghoo Cho, and Christopher Olston. “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. New York: ACM Press, 2004.
- [Page et al. 98] 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.
- [Xue et al. 03] Gui-Rong Xue, Hua-Jun Zeng, Zheng Chen, Wei-Ying Ma, Hong-Jiang Zhang, and Chao-Jun Lu. “Implicit Link Analysis for Small Web Search.” In *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 56–63. New York: ACM Press, 2003.
- [Yu et al. 04] Philip S. Yu, Xin Li, and Bing Liu. “On the Temporal Dimension of Search.” In *Proceedings of the 13th International World Wide Web Conference on Alternate Track Papers & Posters*, pp. 448–449. New York: ACM Press, 2004.

Klaus Berberich, Max Planck Institute for Informatics, Stuhlsatzenhausweg 85, D-66123 Saarbrücken, Germany (kberberi@mpi-inf.mpg.de)

Michalis Vazirgiannis, Department of Informatics, Athens University of Economics & Business, Patision 76, 10434 Athens, Greece (mvazirg@aueb.gr)

Gerhard Weikum, Max Planck Institute for Informatics, Stuhlsatzenhausweg 85, D-66123 Saarbrücken, Germany (weikum@mpi-inf.mpg.de)

Received November 1, 2004; accepted August 31, 2005.