

A Dynamical System for PageRank with Time-Dependent Teleportation

David F. Gleich Ryan A. Rossi

Purdue University

Department of Computer Science

305 N. University St., West Lafayette, IN 47906

{dgleich, rrossi}@purdue.edu

November 20, 2012

We propose a dynamical system that captures changes to the network centrality of nodes as external interest in those nodes vary. We derive this system by adding time-dependent teleportation to the PageRank score. The result is not a single set of importance scores, but rather a time-dependent set. These can be converted into ranked lists in a variety of ways, for instance, by taking the largest change in the importance score. For an interesting class of the dynamic teleportation functions, we derive closed form solutions for the dynamic PageRank vector. The magnitude of the deviation from a static PageRank vector is given by a PageRank problem with complex-valued teleportation parameters. Moreover, these dynamical systems are easy to evaluate. We demonstrate the utility of dynamic teleportation on both the article graph of Wikipedia, where the external interest information is given by the number of hourly visitors to each page, and the Twitter social network, where external interest is the number of tweets per month. For these problems, we show that using information from the dynamical system helps improve a prediction task and identify trends in the data.

1 Introduction

The PageRank vector of a directed graph is the stationary distribution of a Markovian *random surfer*. At a node, the random surfer either

1. transitions to a new node from the set of out-edges, or
2. does something else (e.g. leaves the graph and then randomly returns) [Page et al., 1998, Langville and Meyer, 2006].

The probability that the surfer performs the first action is known as the damping parameter in PageRank, denoted α . The second action is called teleporting and is modeled by the surfer picking a node at random according to a distribution called the teleportation

distribution vector or personalization vector. This PageRank Markov chain always has a unique stationary distribution for any $0 \leq \alpha < 1$. In this paper, we focus on the teleportation distribution vector \mathbf{v} and study how changing teleportation behavior manifests itself in a dynamical system formulation of PageRank.

To proceed further, we need to formalize the PageRank model. Let \mathbf{A} be the adjacency matrix for a graph where $A_{i,j}$ denotes an edge from node i to node j . To avoid a proliferation of transposes, we define \mathbf{P} as the transposed transition matrix for a random-walk on a graph:

$$P_{j,i} = \text{probability of transitioning from node } i \text{ to node } j.$$

Hence, the matrix \mathbf{P} is *column-stochastic* instead of row-stochastic, which is the standard in probability theory. Throughout this manuscript, we utilize uniform random-walks on a graph, in which case $\mathbf{P} = \mathbf{A}^T \mathbf{D}^{-1}$ where \mathbf{D} is a diagonal matrix with the degree of each node on the diagonal. However, none of the theory is restricted to this type of random walk and any column-stochastic matrix will do. If any nodes have no out-links, we assume that they are adjusted in one of the standard ways [Boldi et al., 2007]. Let \mathbf{v} be a teleportation distribution vector such that $v_i \geq 0$ and $\sum_i v_i = 1$. This vector models where the surfer will transition when “doing something else.” The PageRank Markov chain then has the transition matrix:

$$\alpha \mathbf{P} + (1 - \alpha) \mathbf{v} \mathbf{e}^T.$$

While finding the stationary distribution of a Markov chain usually involves computing an eigenvector or solving a singular linear system, the PageRank chain has a particularly simple form for the stationary distribution vector \mathbf{x} :

$$(\mathbf{I} - \alpha \mathbf{P}) \mathbf{x} = (1 - \alpha) \mathbf{v}.$$

Sensitivity of PageRank with respect to \mathbf{v} is fairly well understood. Langville and Meyer [2006] devote a section to determining the Jacobian of the PageRank vector with respect to \mathbf{v} . The choice of \mathbf{v} is often best guided by an application specific measure. By setting $\mathbf{v} = \mathbf{e}_i$, that is, the i th canonical basis vector:

$$\mathbf{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \text{ } i\text{th row,}$$

PageRank computes a highly localized diffusion that is known to produce empirically meaningful clusters and theoretically supported clusters [Andersen et al., 2006, Tong et al., 2006]. By choosing \mathbf{v} based on a set of known-to-be-interesting nodes, PageRank will compute an expanded set of interesting nodes [Gyöngyi et al., 2004, Singh et al.,

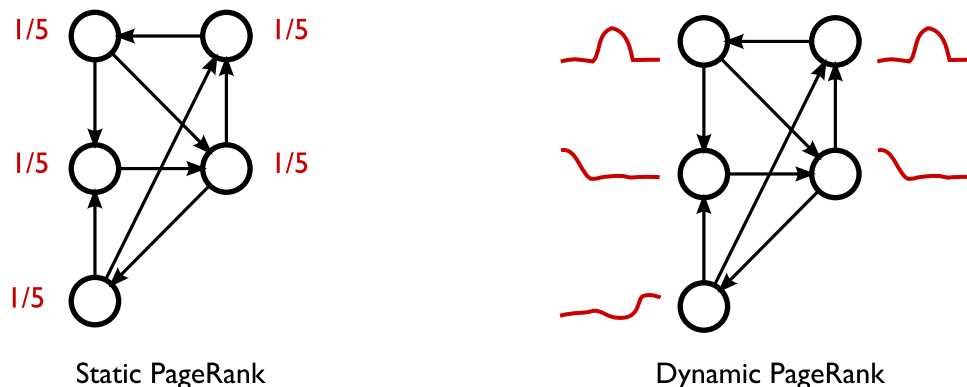


Figure 1 – At left, we have PageRank with static teleportation. At each step, the teleportation is to each node with uniform probability $1/5$. At right, we have the PageRank model with dynamic teleportation. In this case, the teleportation distribution (illustrated in red) *changes with time*. Thus, the upper nodes are teleported to more during the middle time regime. In both cases, the graph is fixed. We study the effect of such dynamic teleportation on the PageRank scores.

2007]. Yet, in all of these cases, \mathbf{v} is chosen once for the graph application or particular problem.

In the original motivation of PageRank [Page et al., 1998], the distribution \mathbf{v} should model how users behave on the web when they don’t click a link. When this intuition is applied to a site like Wikipedia, this suggests that the teleportation function should vary as particular topics become interesting. For instance, in our experiments (section 5), we examine the number of page views for each Wikipedia article during a period where a major earthquake occurred. Suddenly, page views to earthquake spike – presumably as people are searching for that phrase. We wish to *include* this behavior into our PageRank model to understand what is now important in light of a radically different behavior. One option would be to recompute a new PageRank vector given the observed teleporting behavior at the current time. Our proposal for a dynamical system is another alternative. That is, we define a new model where teleportation is the time-dependent function:

$$\mathbf{v}(t).$$

At each time t , $\mathbf{v}(t)$ is a probability distribution of where the random walk teleports. Figure 1 illustrates this model. We return to a comparison between this approach and solving PageRank systems in section 4.

The dynamical system we propose is a generalization of PageRank in the sense that if $\mathbf{v}(t)$ is a constant function in time, then we converge to the standard PageRank vector (theorem 4). Additionally, we can analyze the dynamical PageRank function for some simple oscillatory teleportation functions $\mathbf{v}(t)$. Bounding the deviation of these oscillatory PageRank values from the static PageRank vector involves solving a *PageRank*

problem with complex teleportation [Horn and Serra-Capizzano, 2007, Constantine and Gleich, 2010]. This result is, perhaps, the first non-analytical use of PageRank with a complex teleportation parameter.

In our new dynamical system, we do not compute a single ranking vector as others have done with time-dependent rankings [Grindrod et al., 2011], rather we compute a time-dependent ranking function $\mathbf{x}(t)$, the dynamic PageRank vector at time t , from which we can extract different static rankings (section 2.4). There are two complications with using empirically measured data. First, we must choose a time-scale for our ODE based on the period of our page-view data (section 2.5). Put a bit informally, we must pick the time-unit for our ODE – it is not dimensionless. We analytically show that some choices of the time-scale amount to solving the PageRank system for each change in the teleportation vector. Second, we also investigate smoothing the measured page view data (section 2.6). To compute this dependent ranking function $\mathbf{x}(t)$ we discuss ordinary differential equation (ODE) integrators in section 3.

We discuss the impact of these choices on two problems: page views from Wikipedia and a retweet network from Twitter. We also investigate how the rankings extracted from our methods differ from those extracted by other static ranking measurements. We can use these rankings for a few interesting applications. Adding the dynamic PageRank scores to a prediction task decreases the average error (section 6.1) for Twitter. Clustering the dynamic PageRank scores yields many of the standard time-series features in social networks (section 6.2). Finally, using Granger causality testing on the dynamic PageRank scores helps us find a set of interesting links in the graph (section 6.3).

We make our code and data available in the spirit of reproducible research:

<http://www.cs.purdue.edu/homes/dgleich/codes/dynsyspr-im>

2 PageRank with time-dependent teleportation

We begin our discussion by summarizing the notation introduced thus far in Table 1.

In order to incorporate changes in the teleportation into a new model for PageRank, we begin by reformulating the standard PageRank algorithm in terms of changes to the PageRank values for each page. This step allows us to state PageRank as a dynamical system, in which case we can easily incorporate changes into the vector.

The standard PageRank algorithm is the power method for the PageRank Markov chain [Langville and Meyer, 2006]. After simplifying this iteration by assuming that $\mathbf{e}^T \mathbf{x} = 1$, it becomes:

$$\mathbf{x}^{(k+1)} = \alpha \mathbf{P} \mathbf{x}^{(k)} + (1 - \alpha) \mathbf{v}.$$

In fact, this iteration is equivalent to the Richardson iteration for the PageRank linear system $(\mathbf{I} - \alpha \mathbf{P}) \mathbf{x} = (1 - \alpha) \mathbf{v}$. This fact is relevant because the Richardson iteration is usually defined:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(x)} + \omega \left[(1 - \alpha) \mathbf{v} - (\mathbf{I} - \alpha \mathbf{P}) \mathbf{x}^{(k)} \right].$$

Table 1 – Summary of notation. Matrices are bold, upright roman letters; vectors are bold, lowercase roman letters; and scalars are unbolded roman or greek letters. Indexed elements are vectors if they are bolded, or scalars if unbolded.

i	the imaginary number
n	number of nodes in a graph
\mathbf{e}	the vector of all ones
\mathbf{P}	column stochastic matrix
α	damping parameter in PageRank
\mathbf{v}	teleportation distribution vector
\mathbf{x}	solution to the PageRank computation: $(\mathbf{I} - \alpha\mathbf{P})\mathbf{x} = (1 - \alpha)\mathbf{v}$
$\mathbf{x}(t)$	solution to the dynamic PageRank computation for time t
$\mathbf{v}(t)$	a teleportation distribution vector at time t
\mathbf{c}	the cumulative rank function
\mathbf{r}	the variance rank function
\mathbf{d}	the difference rank function
\mathbf{v}_k	the teleportation distribution for the k th observed page-views vector
θ	decay parameter for time-series smoothing
s	the time-scale of the dynamical system
t_{\max}	the last time of the dynamical system

For $\omega = 1$, we have:

$$\Delta\mathbf{x}^{(k)} = \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} = \alpha\mathbf{P}\mathbf{x}^{(k)} + (1 - \alpha)\mathbf{v} - \mathbf{x}^{(k)} = (1 - \alpha)\mathbf{v} - (\mathbf{I} - \alpha\mathbf{P})\mathbf{x}^{(k)}.$$

Thus, changes in the PageRank values at a node *evolve* based on the value $(1 - \alpha)\mathbf{v} - (\mathbf{I} - \alpha\mathbf{P})\mathbf{x}^{(k)}$. We reinterpret this update as a continuous time dynamical system:

$$\mathbf{x}'(t) = (1 - \alpha)\mathbf{v} - (\mathbf{I} - \alpha\mathbf{P})\mathbf{x}(t). \quad (1)$$

To define the PageRank problem with time-dependent teleportation, we make $\mathbf{v}(t)$ a function of time.

DEFINITION 1 *The dynamic PageRank model with time-dependent teleportation is the solution of*

$$\mathbf{x}'(t) = (1 - \alpha)\mathbf{v}(t) - (\mathbf{I} - \alpha\mathbf{P})\mathbf{x}(t) \quad (2)$$

where $\mathbf{x}(0)$ is a probability distribution vector and $\mathbf{v}(t)$ is a probability distribution vector for all t .

In the dynamic PageRank model, the PageRank values $\mathbf{x}(t)$ may not “settle” or converge to some fixed vector \mathbf{x} . We see this as a feature of the new model as we plan to utilize information from the evolution and changes in the PageRank values. For instance, in [section 2.4](#), we discuss various functions of $\mathbf{x}(t)$ that define a rank. Next, we state the solution of the problem.

LEMMA 2 *The solution of the dynamical system:*

$$\mathbf{x}'(t) = (1 - \alpha)\mathbf{v}(t) - (\mathbf{I} - \alpha\mathbf{P})\mathbf{x}(t)$$

is

$$\mathbf{x}(t) = \exp[-(\mathbf{I} - \alpha\mathbf{P})t]\mathbf{x}(0) + (1 - \alpha) \int_0^t \exp[-(\mathbf{I} - \alpha\mathbf{P})(t - \tau)]\mathbf{v}(\tau) d\tau.$$

This result is found in standard texts on dynamical systems, for example [Berman et al., 1989].

Given this solution, let us quickly verify a few properties of this system:

LEMMA 3 *The solution of a dynamical PageRank system $\mathbf{x}(t)$ is a probability distribution ($\mathbf{x}(t) \geq 0$ and $\mathbf{e}^T \mathbf{x}(t) = 1$) for all t .*

Proof The model requires that $\mathbf{x}(0)$ is a probability distribution. Thus, $\mathbf{x}(0) \geq 0$ and $\mathbf{e}^T \mathbf{x}(0) = 1$. Assuming that the sum of $\mathbf{x}(t)$ is 1, then the sum of the derivative $\mathbf{x}'(t)$ is 0 as a quick calculation shows. The closed form solution above is also nonnegative because the matrix $\exp[-(\mathbf{I} - \alpha\mathbf{P})] = \exp(\alpha\mathbf{P}) \exp(-1) \geq 0$ and both $\mathbf{x}(0)$ and $\mathbf{v}(t)$ are non-negative for all t . (This property is known as exponential non-negativity and it is another property of M -matrices such as $\mathbf{I} - \alpha\mathbf{P}$ [Berman et al., 1989].) ■

2.1 A generalization of PageRank

This closed form solution can be used to solve a version of the dynamic problem that reduces to the PageRank problem with static teleportation. If $\mathbf{v}(t) = \mathbf{v}$ is constant with respect to time, then

$$\int_0^t \exp[-(\mathbf{I} - \alpha\mathbf{P})(t - \tau)]\mathbf{v}(\tau) d\tau = (\mathbf{I} - \alpha\mathbf{P})^{-1}\mathbf{v} - \exp[-(\mathbf{I} - \alpha\mathbf{P})t](\mathbf{I} - \alpha\mathbf{P})^{-1}\mathbf{v}.$$

Hence, for constant $\mathbf{v}(t)$:

$$\mathbf{x}(t) = \exp[-(\mathbf{I} - \alpha\mathbf{P})t](\mathbf{x}(0) - \mathbf{x}) + \mathbf{x},$$

where \mathbf{x} is the solution to static PageRank: $(\mathbf{I} - \alpha\mathbf{P})\mathbf{x} = (1 - \alpha)\mathbf{v}$. Because all the eigenvalues of $-(\mathbf{I} - \alpha\mathbf{P})$ are less than 0, the matrix exponential terms disappear in a sufficiently long time horizon. Thus, when $\mathbf{v}(t) = \mathbf{v}$, nothing has changed. We recover the original PageRank vector \mathbf{x} as the steady-state solution:

$$\lim_{t \rightarrow \infty} \mathbf{x}(t) = \mathbf{x} \text{ the PageRank vector.}$$

This derivation shows that dynamic teleportation PageRank is a generalization of the PageRank vector. We summarize this discussion as:

THEOREM 4 *PageRank with time-dependent teleportation is a generalization of PageRank. If $\mathbf{v}(t) = \mathbf{v}$, then the solution of the ordinary differential equation:*

$$\mathbf{x}'(t) = (1 - \alpha)\mathbf{v} - (\mathbf{I} - \alpha\mathbf{P})\mathbf{x}(t)$$

converges to the PageRank vector

$$(\mathbf{I} - \alpha\mathbf{P})\mathbf{x} = (1 - \alpha)\mathbf{v}$$

as $t \rightarrow \infty$.

2.2 Choosing the initial condition

There are three natural choices for the initial condition $\mathbf{x}(0)$. The first choice is the uniform vector $\mathbf{x}(0) = \frac{1}{n}\mathbf{e}$. The second choice is the initial teleportation vector $\mathbf{x}(0) = \mathbf{v}(0)$. And the third choice is the solution of the PageRank problem for the initial teleportation vector $(\mathbf{I} - \alpha\mathbf{P})\mathbf{x}(0) = (1 - \alpha)\mathbf{v}(0)$. We recommend either of the latter two choices in order to generalize the properties of PageRank. Note that if $\mathbf{x}(0)$ is chosen to solve the PageRank system for $\mathbf{v}(0)$, then $\mathbf{x}(t) = \mathbf{x}$ for all t is the solution of the PageRank dynamical system with constant teleportation (theorem 4).

2.3 PageRank with fluctuating interest

One of the advantages of the PageRank dynamical system is that we can study problems analytically. We now do so with the following teleportation function, or forcing function as it would be called in the dynamical systems literature:

$$\mathbf{v}(t) = \frac{1}{k} \sum_{j=1}^k \mathbf{v}_j \left(\cos\left(t + (j-1)\frac{2\pi}{k}\right) + 1 \right),$$

where \mathbf{v}_j is a teleportation vector. Here, the idea is that \mathbf{v}_j represents the propensity of people to visit certain nodes at different times. To be concrete, we might have \mathbf{v}_1 correspond to news websites that are visited more frequently during the morning, \mathbf{v}_2 correspond to websites visited at work, and \mathbf{v}_3 correspond to websites visited during the evening. This function has all the required properties that we need to be a valid teleportation function. With the risk of being overly formal, we'll state these as a lemma.

LEMMA 5 *Let $k \geq 2$. Let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be probability distribution vectors. The time-dependent teleportation function*

$$\mathbf{v}(t) = \frac{1}{k} \sum_{j=1}^k \mathbf{v}_j \left(\cos\left(t + (j-1)\frac{2\pi}{k}\right) + 1 \right),$$

satisfies the both properties:

1. $\mathbf{v}(t) \geq 0$ for all t , and
2. $\sum_{i=1}^n v(t)_i = 1$ for all t

Proof The first property follows directly because the minimum value of the cosine function is -1 , and thus, $\mathbf{v}(t)$ is always non-negative. The second property is also straightforward. Note that

$$\sum_{i=1}^n v(t)_i = 1 + \sum_{j=1}^k \cos(t + (j-1)\frac{2\pi}{k}) = 1 + \sum_{j=1}^k \operatorname{Re}\{\exp(it + \iota(j-1)\frac{2\pi}{k})\}.$$

Let $r_j(t) = \exp(it + \iota(j-1)\frac{2\pi}{k})$. For $t = 0$, these terms express the k roots of unity. For any other t , we simply rotate these roots. Thus we have $\sum_j r_j(t) = 0$ for any t because the sum of the roots of unity is 0 if $k \geq 2$. The second property now follows because the sum of the real component is still zero. \blacksquare

For this function, we can solve for the steady-state solution analytically.

LEMMA 6 Let $k \geq 2$, $0 \leq \alpha < 1$, \mathbf{P} be column-stochastic, $\mathbf{v}_1, \dots, \mathbf{v}_k$ be probability distribution vectors, and

$$\mathbf{v}(t) = \frac{1}{k} \sum_{j=1}^k \mathbf{v}_j \left(\cos(t + (j-1)\frac{2\pi}{k}) + 1 \right) = \frac{1}{k} \mathbf{V} \cos(t + \mathbf{f}) + \frac{1}{k} \mathbf{V} \mathbf{e},$$

where $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_k]$ and $f_j = (j-1)\frac{2\pi}{k}$ $j = 1, \dots, k$. Then the steady state solution of

$$\mathbf{x}'(t) = (1 - \alpha)\mathbf{v}(t) - (\mathbf{I} - \alpha\mathbf{P})\mathbf{x}(t)$$

is

$$\mathbf{x}(t) = \mathbf{x} + \operatorname{Re}\{\mathbf{s} \exp(it)\}$$

where \mathbf{x} is the solution of the static PageRank problem

$$(\mathbf{I} - \alpha\mathbf{P})\mathbf{x} = (1 - \alpha)\frac{1}{k}\mathbf{V} \mathbf{e}$$

and \mathbf{s} is the solution of the static PageRank problem with complex teleportation

$$(\mathbf{I} - \frac{\alpha}{1+\iota}\mathbf{P})\mathbf{s} = (1 - \alpha)\frac{1}{k(1+\iota)}\mathbf{V} \exp(\iota\mathbf{f}).$$

Proof This proof is mostly a derivation of the expression for the solution by guessing the form. First note that if

$$\mathbf{x}(t) = \mathbf{x} + \mathbf{y}(t)$$

then

$$\mathbf{y}'(t) = (1 - \alpha)\frac{1}{k}\mathbf{V} \cos(t + \mathbf{f}) - (\mathbf{I} - \alpha\mathbf{P})\mathbf{y}(t).$$

That is, we've removed the constant term from the teleportation function by looking at solutions centered around the static PageRank solution. To find the steady-state solution, we look at the complex-phasor problem:

$$\mathbf{z}'(t) = (1 - \alpha)\frac{1}{k}\mathbf{V} \exp(it + i\mathbf{f}) - (\mathbf{I} - \alpha\mathbf{P})\mathbf{z}(t)$$

where $\mathbf{y}(t) = \text{Re}\{\mathbf{z}(t)\}$. Suppose that $\mathbf{z}(t) = \mathbf{s} \exp(it)$. Then:

$$\mathbf{z}'(t) = i\mathbf{s} \exp(it) = (1 - \alpha)\frac{1}{n} \exp(it) \exp(i\mathbf{f}) - (\mathbf{I} - \alpha\mathbf{P})\mathbf{s} \exp(it).$$

The statement of \mathbf{s} in the theorem is exactly the solution after canceling the phasor $\exp(it)$. We now have to show that this solution is well-defined. PageRank with a complex teleportation parameter γ exists for any column-stochastic \mathbf{P} if $|\gamma| < 1$ (see [Horn and Serra-Capizzano \[2007\]](#), [Constantine and Gleich \[2010\]](#)). For the problem defining \mathbf{s} , $\gamma = \alpha/(1 + i)$ and $|\gamma| = \alpha/\sqrt{2}$. Thus, such a vector \mathbf{s} always exists. ■

We conclude with an example of this theorem. Consider a four node graph with adjacency matrix and transition matrix:

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{P} = \begin{bmatrix} 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0.5 & 0 \end{bmatrix}.$$

Let $\mathbf{v}_j = \mathbf{e}_j$ for $j = 1, \dots, 4$. That is, interest oscillates between all four nodes in the graph in a regular fashion. We show the evolution of the dynamical system for 20 time-units in [Figure 2](#). This evolution quickly converges to the oscillators predicted by the lemma. In the interest of simplifying the plot, we do not show the exact curves as they are visually indistinguishable from those plotted for $t \geq 4$. By solving the complex valued PageRank to compute \mathbf{s} , we can compute the magnitude of the fluctuation:

$$|\mathbf{s}| = \begin{bmatrix} 0.0216 & 0.0261 & 0.0122 & 0.0235 \end{bmatrix}^T.$$

This vector accurately captures the magnitude of these fluctuations.

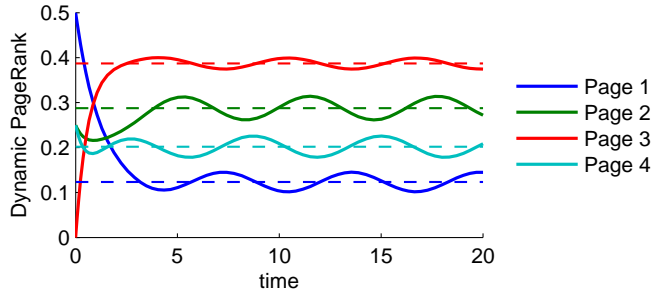


Figure 2 – The dashed lines represent the average PageRank vector computed for the teleportation vectors. The curves show the evolution of the PageRank dynamic system for this example of teleportation. We see that the dynamic PageRanks fluctuate about their average PageRank vectors. [Lemma 6](#) predicts the magnitude of the fluctuation.

2.4 Ranking from Time-Series

The above equations provide a time-series of dynamic PageRank vectors for the nodes, denoted formally as $\mathbf{x}(t)$, $0 \leq t \leq t_{\max}$. Applications, however, often want a single score, or small set of scores, to characterize sets of interesting nodes. There are a few ways in which these time series give rise to scores. Many of these methods were explained by O'Madadhain and Smyth [2005] in the context of ranking sequences of vectors. Having a variety of different scores derived from the same data frequently helps when using these scores as features in a prediction or learning task [Becchetti et al., 2008, Constantine and Gleich, 2010].

Transient Rank. We call the instantaneous values of $\mathbf{x}(t)$ a node's *transient* rank. This score gives the importance of a node at a particular time.

Summary, Variance, & Cumulative Rank. Any summary function s of the time series, such as the integral, average, minimum, maximum, or variance, is a single score that encompasses the entire interval $[0, t_{\max}]$. We utilize the *cumulative rank*, \mathbf{c} and *variance rank*, \mathbf{r} in the forthcoming experiments:

$$\mathbf{c} = \int_0^{t_{\max}} \mathbf{x}(t) dt \quad \text{and} \quad \mathbf{r} = \int_0^{t_{\max}} \left(\mathbf{x}(t) - \frac{1}{t_{\max}} \mathbf{c} \right)^2 dt.$$

Difference Rank. A node's difference rank is the difference between its maximum and minimum rank over all time, or a limited time window:

$$\mathbf{d} = \max_t [\mathbf{x}(t)] - \min_t [\mathbf{x}(t)] \quad \mathbf{d}_W = \max_{t \in W} \mathbf{x}(t) - \min_{t \in W} \mathbf{x}(t).$$

Nodes with high difference rank should reflect important events that occurred within the range $[0, t_{\max}]$ or the time window W . We suggest using a window W that omits the initial convergence region of the evolution. In the context of Figure 2, we'd set W to be $[4, 20]$ to approximate the vector $|\mathbf{s}|$ numerically. In Section 6 and figure 7, we see examples of how current news stories arise as articles with high difference rank.

2.5 Modeling activity

In the next two sections of our introduction to the dynamic teleportation PageRank model, we discuss how to incorporate empirically measured activity into the model. Let $\mathbf{p}_1, \dots, \mathbf{p}_k$ be k observed vectors of activity for a website. In the cases we examine below, these activity vectors measure page views per hour on Wikipedia and the number of tweets per month on Twitter. We normalize each of them into teleportation distributions, and conceptually think of the collection of vectors as a matrix

$$\mathbf{v}_1, \dots, \mathbf{v}_k \quad \rightarrow \quad \mathbf{V} = \left[\mathbf{v}_1, \dots, \mathbf{v}_k \right].$$

Let $\mathbf{e}(i)$ be a functional form representing the vector \mathbf{e}_i . The time-dependent teleportation vector we create from this data is:

$$\mathbf{v}(t) = \mathbf{V}\mathbf{e}(\text{floor}\{t\} + 1) = \mathbf{v}_{\text{floor}\{t\}+1}.$$

For this choice, the time-units of our dynamical system are given by the time-unit of the original measurements. Other choices are possible too. Consider:

$$\mathbf{v}_s(t) = \mathbf{V}\mathbf{e}(\text{floor}\{t/s\} + 1) = \mathbf{v}_{\text{floor}\{t/s\}+1}.$$

If $s > 1$, then time in the dynamical system slows down. If $s < 1$, then time accelerates. Thus, we call s the time-scale of the system. Note that

$$\mathbf{x}(sj), \quad j = 0, 1, \dots$$

represents the same effective time-point for any time-scale. Thus, when we wish to compare different time-scales s , we examine the solution at such scaled points.

In the experimental evaluation, the parameter s plays an important role. We illustrate its effect in Figure 3(a) for a small subnetwork extracted from Wikipedia. As we discuss further in section 3, for large values of s , then $\mathbf{v}(t)$ looks constant for long periods of time, and hence $\mathbf{x}(t)$ begins to converge to the PageRank vector for the current, and effectively static, teleportation vector. Thus, we also plot the converged PageRank vectors as a step function. We see that as s increases, the lines converge to these step functions, but for $s = 1$ and $s = 2$, they behave differently.

2.6 Smoothing empirical activity

So far, we defined a time-dependent that $\mathbf{v}(t)$ changes at fixed intervals based on empirically measured data. A better idea is to smooth out these “jumps” using an exponentially weighted moving average. As a continuous time function, this yields:

$$\bar{\mathbf{v}}'(t; \theta) = \theta\mathbf{v}(t) - \theta\bar{\mathbf{v}}(t; \theta).$$

To understand why this smooths the sequence, consider an implicit Euler approximation:

$$\bar{\mathbf{v}}(t) = \frac{1}{1+h\theta}\bar{\mathbf{v}}(t-h; \theta) + \frac{h\theta}{1+h\theta}\mathbf{v}(t).$$

This update can be written more simply as:

$$\bar{\mathbf{v}}(t; \theta) = \underbrace{\gamma\mathbf{v}(t)}_{\text{new data}} + \underbrace{(1-\gamma)\bar{\mathbf{v}}(t-h; \theta)}_{\text{old data}},$$

where $\gamma = \frac{h\theta}{1+h\theta}$. When $\mathbf{v}(t)$ changes at fixed intervals, then $\bar{\mathbf{v}}(t; \theta)$ will slowly change. If θ is small then $\bar{\mathbf{v}}(t; \theta)$ changes slowly. We recover the “jump” changes in $\mathbf{v}(t)$ in the limit $\theta \rightarrow \infty$.

The effect of θ is shown in Figure 3(b). Note that we quickly recover behavior that is effectively the same as using jumps in $\mathbf{v}(t)$ ($\theta = 1, 10$). So we only expect changes with smoothing for $\theta < 1$.

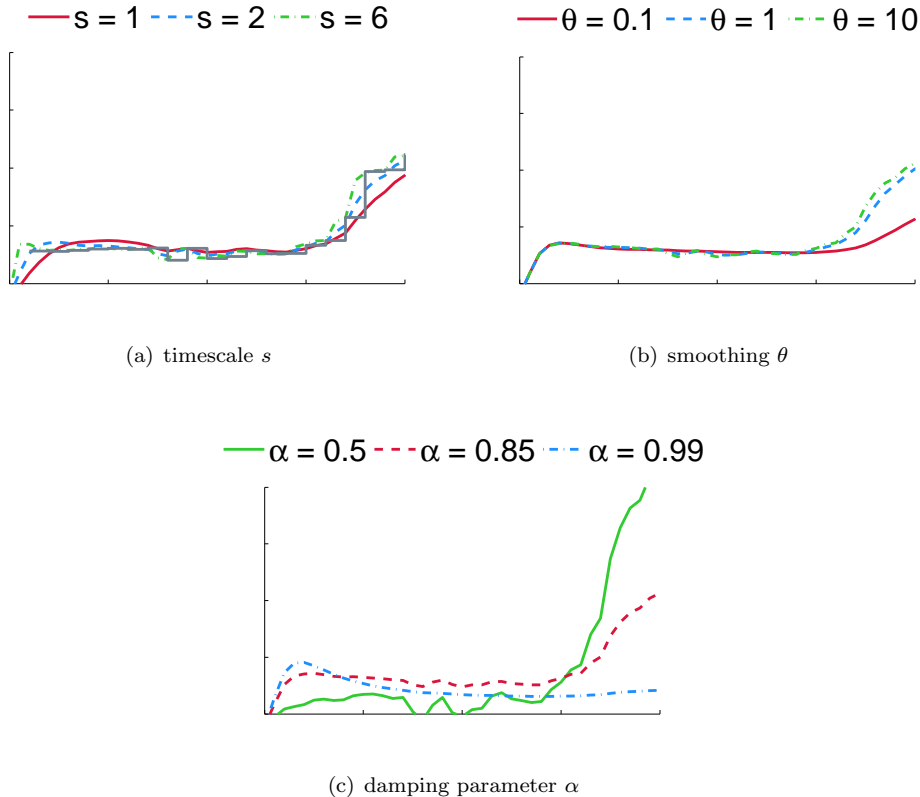


Figure 3 – The evolution of PageRank values for one node due to the dynamical teleportation. The horizontal axis is time $[0, 20]$, and the vertical axis runs between $[0.01, 0.014]$. In figure (a), $\alpha = 0.85$, and we vary the time-scale parameter (section 2.5) with no smoothing. The solid dark line corresponds to the step function of solving PageRank exactly at each change in the teleportation vector. All samples are taken from the same effective time-points as discussed in the section. In figure (b), we vary the smoothing (section 2.6) of the teleportation vectors with $s = 2$, and $\alpha = 0.85$. In figure (c), we vary α with $s = 2$ and no smoothing. We used the `ode45` function in Matlab, a Runge-Kutta method, to evolve the system.

2.7 Choosing the teleportation factor

Picking α even for static PageRank problems is challenging, see Gleich et al. [2010] and Constantine and Gleich [2010] for some discussion. In this manuscript, we do not perform any systematic study of the effects of α beyond Figure 3(c). This simple experiment shows one surprising feature. Common wisdom for choosing α in the static case suggests that as α approaches 1, the vector becomes more sensitive. For the dynamic teleportation setting, however, the opposite is true. Small values of α produce solutions that more closely reflect the teleportation vector – the quantity that is changing – whereas

large values of α reflect the graph structure, which is invariant with time. Hence, with dynamic teleportation, using a small value of α is the sensitive setting. Note that this observation is a straightforward conclusion from the equations of the dynamic vector:

$$\mathbf{x}'(t) = (1 - \alpha)\mathbf{v}(t) + \alpha\mathbf{P}\mathbf{x}(t) - \mathbf{x}(t)$$

so α small implies a larger change due to $\mathbf{v}(t)$. Nevertheless, we found it surprising in light of the existing literature.

3 Methods for dynamic PageRank

In order to compute the time-sequence of PageRank values $\mathbf{x}(t)$, we can evolve the dynamical system (2) using any standard method – usually called an integrator. We discuss both the forward Euler method and a Runge-Kutta method next. Both methods, and indeed, the vast majority of dynamical system integrators only require a means to evaluate the derivative of the system at a time t given $\mathbf{x}(t)$. For PageRank with dynamic teleportation, this corresponds to computing:

$$\mathbf{x}'(t) = (1 - \alpha)\mathbf{v}(t) - (\mathbf{I} - \alpha\mathbf{P})\mathbf{x}(t).$$

The dominant cost in evaluating $\mathbf{x}'(t)$ is the matrix vector product $\mathbf{P}\mathbf{x}$. For the explicit methods we explore, all of the other work is linear in the number of nodes, and hence, these methods easily scale to large networks. Both of these methods may also be used in a distributed setting if a distributed matrix-vector product is available.

3.1 Forward Euler

We first discuss the forward Euler method. This method lacks high accuracy, but is fast and straightforward. Forward Euler approximates the derivative with a first order Taylor approximation:

$$\mathbf{x}'(t) \approx \frac{\mathbf{x}(t+h) - \mathbf{x}(t)}{h},$$

and then uses that approximation to estimate the value at a short time-step in the future:

$$\mathbf{x}(t+h) = \mathbf{x}(t) + h[(1 - \alpha)\mathbf{v}(t) - (\mathbf{I} - \alpha\mathbf{P})\mathbf{x}(t)].$$

This update is the original Richardson iteration with $h = \omega$. We present the forward Euler method as a formal algorithm in Figure 4 in order to highlight a comparison with the power and Richardson method. That is, the forward Euler method is simply running a power method, but changing the vector \mathbf{v} at every iteration. However, we derived this method based on evolving (2). Thus, by studying the relationship between (2) and the algorithm in Figure 4, we can understand the underlying problem solved by changing the teleportation vector while running the power method.

Input:

a graph $G = (V, E)$ and a procedure to compute $\mathbf{P}\mathbf{x}$ for this graph
 a maximum time t_{\max}
 a function to return $\mathbf{v}(t)$ for any $0 \leq t \leq t_{\max}$
 a damping parameter α
 a time-step h

Output: \mathbf{X} where the k th column of \mathbf{X} is $\mathbf{x}(0 + kh)$ for all $1 \leq k \leq t_{\max}/h$ (or any desired subset of these values)

```

t ← 0; k = 1
x(0) ← v(0) (or any other desired initial condition)
while t ≤ t_max - h do
  x(t + h) ← x(t) + h [(1 - α)v(t) - (I - αP)x(t)]
  X(:, k) ← x(t + h)
  t ← t + h; k ← k + 1

```

Figure 4 – The forward Euler method for evolving the dynamical system: $\mathbf{x}'(t) = (1 - \alpha)\mathbf{v}(t) - (\mathbf{I} - \alpha\mathbf{P})\mathbf{x}(t)$. The resulting procedure looks remarkably similar to the standard Richardson iteration to compute a PageRank vector. One key difference is that there is no notion of convergence.

Long time-scales. Using the forward Euler method, we can analyze the situation with a large time-scale parameter s . Consider an arbitrary $\mathbf{x}(0)$, $\alpha = 0.85$, $s = 100$, $h = 1$, and no smoothing. In this case, then the forward Euler method will run the Richardson iteration for 100 times before observing the change in $\mathbf{v}(t)$ at $t = 100$. The difference between $\mathbf{x}(k)$ and the exact PageRank solution for this temporarily static $\mathbf{v}(t)$ is $\|\mathbf{x}(k) - \mathbf{x}\|_1 \leq 2\alpha^k$. For $k > 50$, this difference is small. Thus, a large s and no smoothing corresponds to solving the PageRank problem for each change in \mathbf{v} .

Stability. The forward Euler method with timestep h is stable if the eigenvalues of the matrix $-h(\mathbf{I} - \alpha\mathbf{P})$ are within distance 1 of the point -1 . The eigenvalues of \mathbf{P} are all between -1 and 1 because it is a stochastic matrix, and so this is stable for any $h < \frac{2}{1+\alpha}$.

3.2 Runge-Kutta

Runge-Kutta [Runge, 1895, Kutta, 1901] numerical schemes are some of the most well-known and most used. They achieve far greater accuracy than the simple forward Euler method, at the expense of a greater number of evaluations of the function $\mathbf{x}'(t)$ at each step. We use the implementations of Runge-Kutta methods available in the Matlab ODE suite [Shampine and Reichelt, 1997]. The step-size is adapted automatically based on a local error estimate, and the solution can be evaluated at any desired point in time. The stability region for Runge-Kutta includes the region for forward Euler, so these methods

are stable. These methods are also fast. To integrate the system for Wikipedia with over 4 million vertices and 60 million edges, it took between 300-600 seconds, depending on the parameters.

3.3 Maintaining interpretability

Based on the theory of the dynamic teleportation system, we expect that $\mathbf{x}(t) \geq 0$ and $\mathbf{e}^T \mathbf{x}(t) = 1$ for all time. Although this property should be true of the computed solution, we often find that the sum diverges from one. Consequently, for our experiments, we include a correcting term:

$$\mathbf{x}'(t) = (1 - \alpha)\mathbf{v}(t) - (\gamma\mathbf{I} - \alpha\mathbf{P})\mathbf{x}(t)$$

where $\gamma = (1 - \alpha)\mathbf{e}^T \mathbf{v}(t) + \alpha\mathbf{e}^T \mathbf{x}(t)$. Note that $\gamma = 1$ if the $\mathbf{x}(t)$ has sum exactly 1. If $\mathbf{e}^T \mathbf{x}(t)$ is slightly different from one, then the correction with γ ensures that $\mathbf{e}^T \mathbf{x}'(t) = 0$ numerically. Similar issues arise in computing static PageRank [Wills and Ipsen, 2009], although the additional computation in the Runge-Kutta methods exacerbates the problem.

4 Related work

Note that we previously studied this idea in a conference paper [Rossi and Gleich, 2012]. These ideas have been significantly refined for this manuscript.

The relationship between dynamical systems and classical iterative methods has been utilized by Embree and Lehoucq [2009] to study eigenvalue solvers. It was also noted in an early paper by Tsaparas [2004] that there is a relationship between the PageRank and HITS algorithms and dynamical systems.

In the past, others studied PageRank approximations on graph streams [Das Sarma et al., 2008]. More recently, Bahmani et al. [2012] studied how accurately an evolving PageRank method could estimate the true PageRank of an evolving graph that is accessed only via a crawler. The method used here solved each PageRank problem exactly for the current estimate of the underlying graph. A detailed study of how PageRank values evolve during a web-crawl was done by Boldi et al. [2005]. In the future, we plan to study dynamic graphs via similar ideas.

As explained in section 3 and figure 4, our proposed method is related to changing the teleportation vector in the power method as its being computed. Bianchini et al. [Bianchini et al., 2005] noted that the power method would still converge if either the graph or the vector \mathbf{v} changed during the method, albeit to a new solution given by the new vector or graph. Our method capitalizes on a closely related idea and we utilize the intermediate quantities explicitly. Another related idea is the Online Page Importance

Computation (OPIC) [Abiteboul et al., 2003], which integrates a PageRank-like computation *with* a crawling process. The method does nothing special if a node has changed when it is crawled again.

While we described PageRank in terms of a random-surfer model, another characterization of PageRank is that it is a sum of damped transitions:

$$\mathbf{x} = (1 - \alpha) \sum_{k=0}^{\infty} (\alpha \mathbf{P})^k \mathbf{v}.$$

These transitions are a type of probabilistic walk and Grindrod et al. [Grindrod et al., 2011] introduced the related notion of dynamic walks for dynamic graphs. We can interpret these dynamic walks as a *backward Euler* approximation to the dynamical system:

$$\mathbf{x}'(t) = \alpha \mathbf{A}(t) \mathbf{x}(t) \quad \mathbf{x}(0) = \mathbf{e}$$

with time-step $h = 1$ and \mathbf{A} is a time-dependent adjacency matrix. This relationship suggests that there may be a range of interesting models between our dynamic teleportation model and existing evolving graph models.

Outside of the context of web-ranking, O'Madadhain and Smyth propose EvenRank [O'Madadhain and Smyth, 2005], a method of ranking nodes in dynamic graphs, that uses the PageRank propagation equations for a sequence of graphs. We utilize the same idea but place it within the context of a continuous dynamical system. In the context of popularity dynamics [Ratkiewicz et al., 2010], our method captures how changes in external interest influence the popularity of nodes and the nodes linked to these nodes in an implicit fashion. Our work is also related to modeling human dynamics, namely, how humans change their behavior when exposed to rapidly changing or unfamiliar conditions [Bagrow et al., 2011]. In one instance, our method shows the important topics and ideas relevant to humans before and after one of the largest Australian Earthquakes (figure 7).

In closing, we wish to note that our proposed method *does not involve* updating the PageRank vector, a related problem which has received considerable attention [Chien et al., 2004, Langville and Meyer, 2004]. Nor is it related to tensor methods for dynamic graph data [Sun et al., 2006, Dunlavy et al., 2011].

5 Examples of dynamic teleportation

We now use dynamic teleportation to investigate page view patterns on Wikipedia and user activity on Twitter. In the following experiments, unless otherwise noted, we set $s = 1$, $\alpha = 0.85$, do not use smoothing (" $\theta = \infty$ "), and use the `ode45` method from Matlab to evolve the system. We study this model on two datasets.

5.1 Datasets

We provide some basic statistics of the Wikipedia and Twitter datasets in Table 2. For Wikipedia, the time unit for $s = 1$ is an hour, and for Twitter, it is one month.

Wikipedia Article Graph and Hourly page views. Wikipedia provides access to copies of its database [Various, 2009]. We downloaded a copy of its database on March 6th, 2009 and extracted an article-by-article link graph, where an article is a page in the main Wikipedia namespace, a category page, or a portal page. All other pages and links were removed. See Gleich et al. [2007] for more information.

Wikipedia also provides hourly page views for each page [Various, 2011]. These are the number of times a page was viewed for a given hour. These are not unique visits. We downloaded the raw page counts and matched the corresponding page counts to the pages in the Wikipedia graph. We used the page counts starting from March 6, 2009 and moving forward in time. Although it would seem like measuring page views would correspond to measuring $\mathbf{x}(t)$ instead of $\mathbf{v}(t)$, one of our earlier studies showed that users hardly ever follow links on Wikipedia [Gleich et al., 2010]. Thus, we can interpret these page views as a reasonable measure of external interest in Wikipedia pages.

Twitter Social Network and Monthly Tweet Rates. We use a follower graph generated by starting with a few seed users and crawling follows links from 2008. We extract the user tweets over time from 2008–2009. A tweet is represented as a tuple (user, time, tweet). Using the set of tweets, we construct a sequence of vectors to represent the number of tweets for a given month.

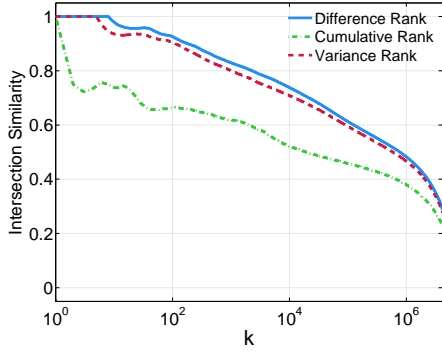
Table 2 – Dataset Properties. The page views or tweets is denoted as \mathbf{p} .

Dataset	Nodes	Edges	t_{\max}	Period	Average p_i	Max p_i
WIKIPEDIA	4,143,840	72,718,664	48	hours	1.4243	353,799
TWITTER	465,022	835,424	6	months	0.5569	1056

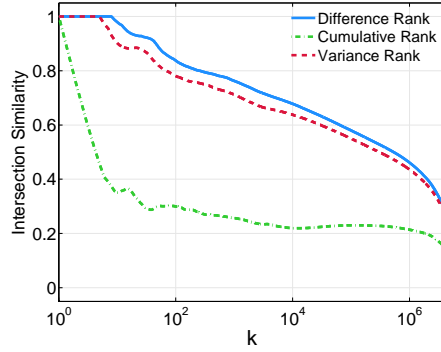
5.2 Rankings from transient scores

First, we evaluate the rankings from dynamic PageRank using the intersection similarity measure [Boldi, 2005]. Given two vectors \mathbf{x} and \mathbf{y} , the intersection similarity metric at k is the average symmetric difference over the top- j sets for each $j \leq k$. If \mathcal{X}_k and \mathcal{Y}_k are the top- k sets for \mathbf{x} and \mathbf{y} , then

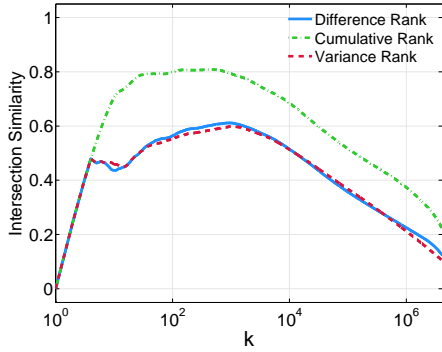
$$\text{isim}_k(\mathbf{x}, \mathbf{y}) = \frac{1}{k} \sum_{j=1}^k \frac{|\mathcal{X}_j \Delta \mathcal{Y}_j|}{2j}$$



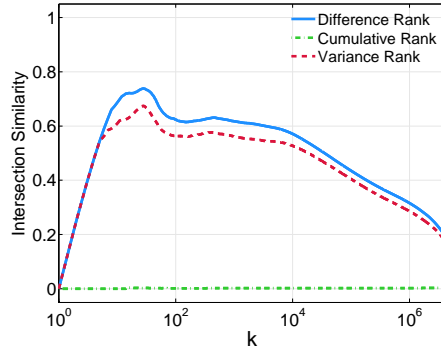
(a) In-degree



(b) Static PageRank (Uniform)



(c) Average page views



(d) Static PageRank (avg. page views)

Figure 5 – Intersection similarity of rankings derived from dynamic PageRank. We compute the intersection similarity of the difference, variance, and cumulative rankings given by dynamic PageRank and compare these with the rankings given by the in-degree, average page views, static PageRank with uniform teleportation, and static PageRank with average page views as the teleportation vector. For dynamic PageRank, we set the initial value $\mathbf{x}(0)$ to be the solution of the static PageRank system which uses $\mathbf{v}(0)$ as the teleportation vector.

where Δ is the symmetric set-difference operation. Identical vectors have an intersection similarity of 0.

For the Wikipedia graph, Figure 5 shows the similarity profile comparing a few ranking measures from dynamic PageRank to reasonable baselines. In particular, we compare \mathbf{d} , \mathbf{r} , \mathbf{c} (from §2.4) to indegree, average page views, static PageRank with uniform teleportation, and static PageRank using average page views as the teleportation vector. The results suggest that dynamic PageRank is different from the other measures, even for small values of k . In particular, combining the external influence with the graph appears to produce something new. The only exception is in Fig. 5(d) where the cumulative rank

is shown to give a similar ordering to static PageRank using average page views as the teleportation.

5.3 Difference ranks

Figure 6 and figure 7 show the time-series of the top 100 pages by the difference measure for Wikipedia with $s = 1$ and $s = 4$ without smoothing. Many of these pages reveal the ability of dynamic PageRank to mesh the network structure with changes in external interest. For instance, in figure 7, we find pages related to an Australian earthquake (43, 84, 82), the “recently” released movie “Watchmen” (98, 23-24), a famous musician that died (2, 75), recent “American Idol” gossip (34, 63), a remembrance of Eve Carson from a contestant on “American Idol” (88, 96, 34), news about the murder of a Harry Potter actor (60), and the Skittles social media mishap (94). These results demonstrate the effectiveness of the dynamic PageRank to identify interesting pages that pertain to external interest. The influence of the graph results in the promotion of pages such as Richter magnitude (84). That page was not in the top 200 from page views.

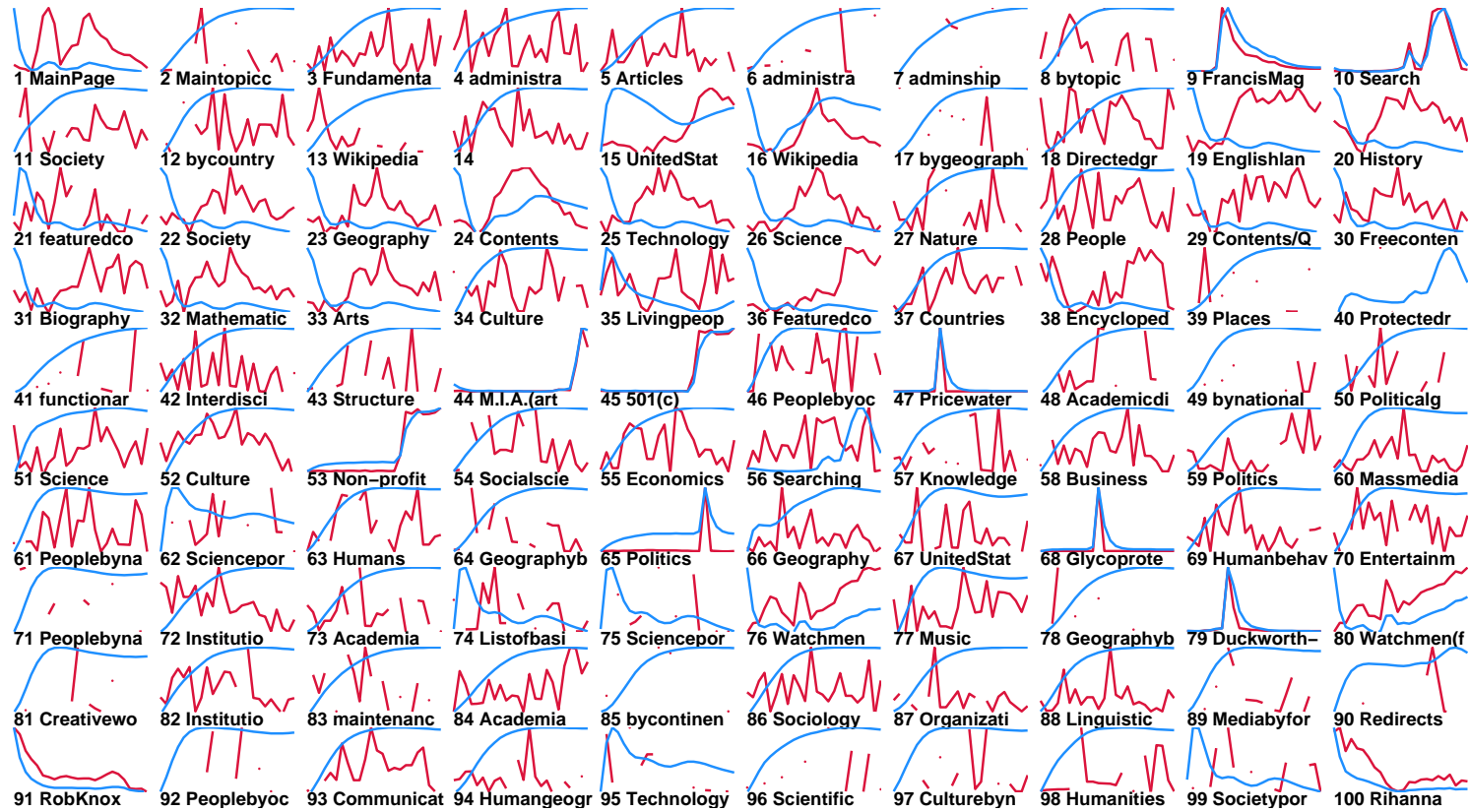


Figure 6 – The top-100 Wikipedia pages that fluctuate the most as determined by the difference ranking from our dynamic PageRank approach. The blue curves are the transient scores, and the red-curves are the raw page view data. The horizontal axis is 24 hours, and the vertical axes are normalized to show the range of the data.

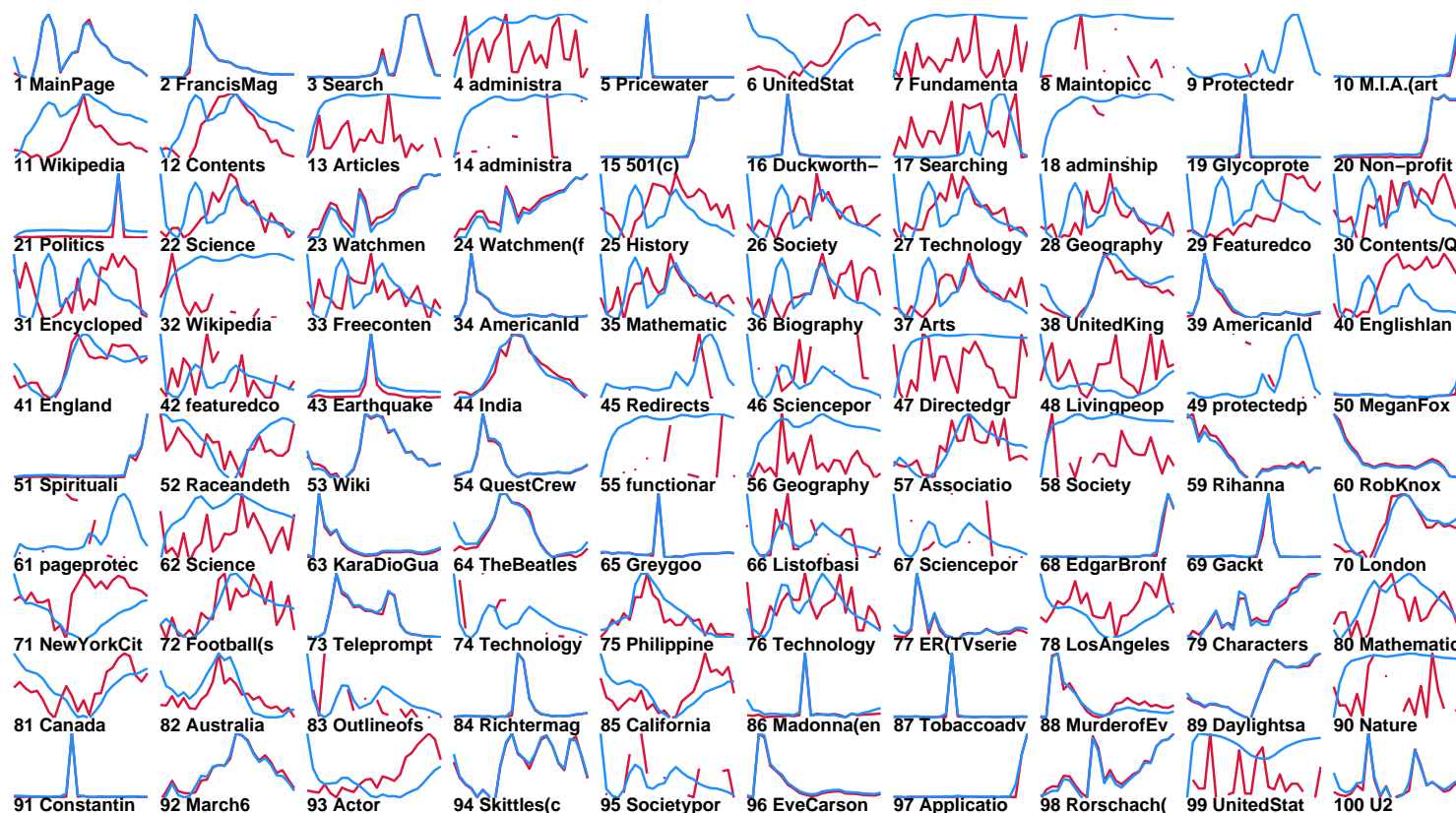


Figure 7 – We again plot the top 100 pages from the difference rank, now using $s = 4$. The blue curves are the transient scores, and the red-curves are the raw page view data. The horizontal axis is 24 hours, and the vertical axes are normalized to show the range of the data. This choice gives a similar ordering as our previous forward Euler iteration method from Rossi and Gleich [2012]. Note the large change between the transient scores for "MainPage" between this figure and figure 6.

6 Applications of time-dependent teleportation

This section explores the opportunity of using Dynamic PageRank for a variety of applications outside of the context of ranking.

6.1 Predicting future page views & tweets

We begin by studying how well the dynamical system can *predict* the future. Formally, given a lagged time-series $\mathbf{p}_{t-w}, \dots, \mathbf{p}_{t-1}, \mathbf{p}_t$ [Ahmed et al., 2010], the goal is to predict the future value \mathbf{p}_{t+1} (actual page views or number of tweets). This type of temporal prediction task has many applications, such as actively adapting caches in large database systems, or dynamically recommending pages.

We performed one-step ahead predictions ($t + 1$) using linear regression. That is, we learn a model of the form:

$$\begin{bmatrix} \bar{\mathbf{f}}(t-1) & \bar{\mathbf{f}}(t-2) & \dots & \bar{\mathbf{f}}(t-w) \end{bmatrix} \mathbf{b} \approx \mathbf{p}(t)$$

where w is the window-size, and $\bar{\mathbf{f}}(\cdot)$ is either page views or both page views and transient scores. After fitting \mathbf{b} , the model predicts $\mathbf{p}(t + 1)$ as

$$\begin{bmatrix} \bar{\mathbf{f}}(t) & \bar{\mathbf{f}}(t-1) & \dots & \bar{\mathbf{f}}(t-w+1) \end{bmatrix} \mathbf{b}$$

We use the symmetric Mean Absolute Percentage Error (sMAPE) [Ahmed et al., 2010] measure to evaluate the prediction:

$$\text{sMAPE} = \frac{1}{|T|} \sum_{t=1}^{|T|} \frac{|p_t - \hat{p}_t|}{(p_t + \hat{p}_t)/2}.$$

This relative error measure averages all the relative prediction errors over all the time-steps. We then average it over nodes.

We study two predictive modes. The *base model* uses only the time-series of page views or tweets to predict the future page views or tweets. The *dynamic teleportation model* uses both the transient scores with smooting and page views to predict the future page views (or tweets).

We evaluate these models for prediction on *stationary* and *non-stationary* time-series. Informally, a time-series is weakly stationary if it has properties (mean and covariance) similar to that of the time-shifted time-series. We consider the top and bottom 10,000 nodes from the difference ranking as nodes that are approximately non-stationary (volatile) and stationary (stable), respectively. Table 3 compares the predictions of the models across time for non-stationary and stationary prediction tasks. Our findings indicate that the Dynamic PageRank time-series provides valuable information for forecasting future tweet rates; however, it adds little (if any) accuracy in forecasting future page views on Wikipedia.

Table 3 – The ratio between the base model and the model with dynamic teleportation scores with $s = 1, 2, 6$, and ∞ , for three smoothing parameters. (Here, $s = \infty$ corresponds to solving the PageRank problem exactly for each change in teleportation.) If this ratio is less than 1, then the model with the dynamic teleportation scores improves the prediction performance. We also distinguish between prediction problems with highly volatile nodes (non-stationary) and nodes with relatively stable behavior (stationary). The results show a much stronger benefit for Twitter than for Wikipedia

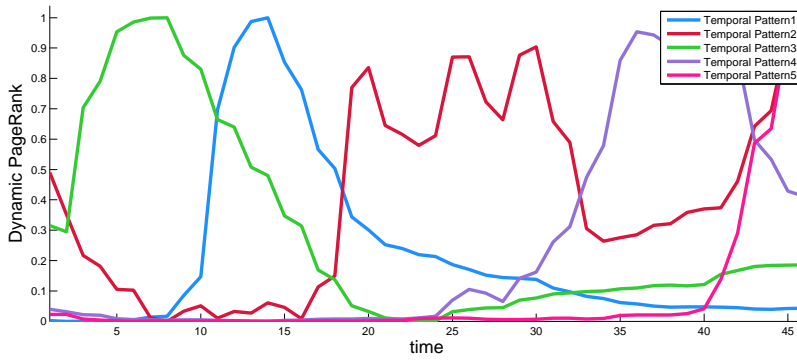
Dataset	Type	θ	Error Ratio			
			s (timescale)			
			1	2	6	∞
TWITTER	stationary	0.01	0.635	0.929	0.913	0.996
		0.50	0.636	0.735	0.854	0.939
		1.00	0.522	0.562	0.710	0.963
	non-stationary	0.01	0.461	0.841	1.001	0.992
		0.50	0.261	0.608	0.585	0.929
		1.00	0.137	0.605	0.617	0.918
WIKIPEDIA	stationary	0.01	0.978	0.991	0.989	0.978
		0.50	1.140	1.130	1.004	0.990
		1.00	1.084	0.976	1.010	0.990
	non-stationary	0.01	0.968	1.011	0.968	1.004
		0.50	1.218	0.994	1.030	1.031
		1.00	1.241	0.996	0.957	0.998

For Twitter, the dynamic teleportation model improves predictions the most with the non-stationary nodes. The diffusion of activity captured by the model allows our model to detect, early on, when the external interest of vertices will change, before that change becomes apparent in the external interest of the vertices. This is easiest to detect when there is a large sudden change in external interest of a neighboring vertex.

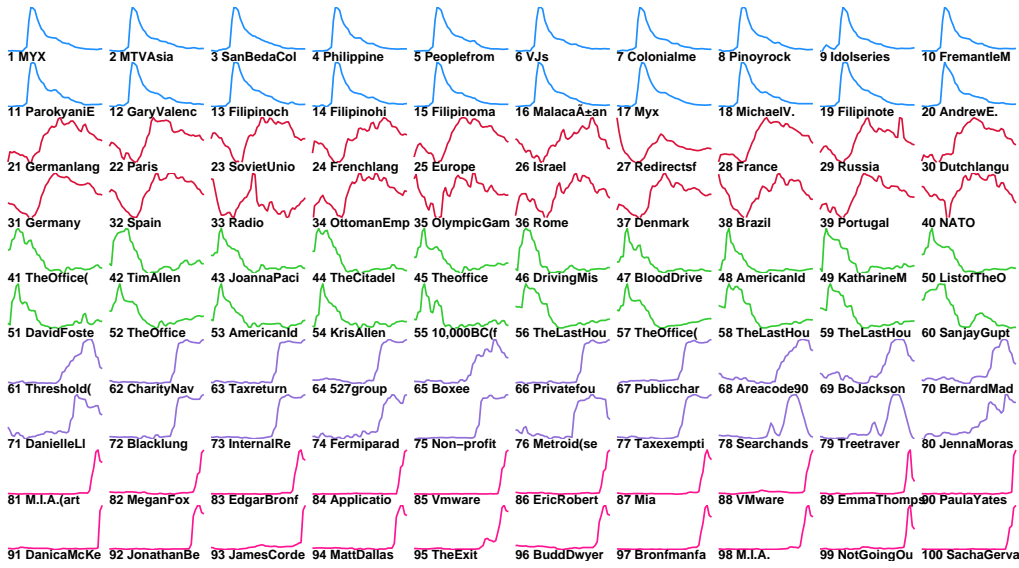
6.2 Clustering transient score trends

Identifying vertices with similar time-series is important for modeling and understanding large collections of multivariate time-series. We now group vertices according to their transient scores. By using the difference rank measure \mathbf{d} for $s = 4$, we cluster the top 5,000 vertices using k-means with $k = 5$, repeat the clustering 2,000 times, and take the minimum distance clustering identified.

The cluster centroids are temporal patterns, and the main patterns in the dynamic PageRanks are visualized in [figure 8\(a\)](#). Pattern 2 represents European-centric behavior, whereas the others correspond to spikes or unusual events occurring within the dynamic PageRank system. [Figure 8\(b\)](#) plots the 20 closest vertices matching the patterns above. A few pages from the five groups are consistent with our previously discussed results from [figure 7](#). One such unusual event is related to the death of a famous musician/actor from the Philippines (see pages 1-20). The pages from the third cluster (41-60) are related to “American Idol” and other TV shows/actors. Also some of the pages from the fourth cluster relate to Bernard Madoff (63, 66, 67, 70, 73), six days before he plead



(a) Temporal Patterns



(b) Vertices with Similar Dynamics

Figure 8 – Vertices with similar dynamical properties are grouped together. The visualization reveals the important dynamic patterns (spikes, trends) present from March 6th, 2009 in our large collection of time-series from Wikipedia. For each hour, we sample twice from the continuous function $\mathbf{x}(t)$ and utilize these intermediate values in the clustering.

guilty in the largest financial fraud in U.S. history. This grouping reveals many of the standard patterns in time-series such as spikes and increasing/decreasing trends [Yang and Leskovec, 2011].

6.3 Towards causal link relationships

In this section, we use Granger causality tests [Granger, 1969] on the collection of transient scores to attempt to understand which links are most important. The Granger causality model, briefly described below, ought to identify a causal relationship between the time-series of any two vertices connected by a directed edge. This is because *there is a causal relationship* between their time-series in our dynamical system. However, due to the impact of the time-dependent teleportation, only some of these links will be identified as causal. We wish to investigate this smaller subset of links.

Intuitively, if a time-series X causally affects another Y , then the past values of X should be helpful in predicting the future values of Y , above what can be predicted based on the past values of Y alone. This is formalized as follows: the error in predicting \hat{y}_{t+s} from y_t, y_{t-1}, \dots should be larger than the error in predicting \hat{y}_{t+s} from the joint data $y_t, y_{t_1}, \dots, x_t, x_{t-1}, \dots$ if X causes Y . As our model, we chose to use the standard vector-autoregressive (VAR) model from econometrics [Box et al., 2011]. This is implemented in a Matlab code by LeSage [1999]. The standard p -lag VAR model takes the form:

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \mathbf{c} + \sum_{i=1}^p \mathbf{M}_i \begin{bmatrix} y_{t-i} \\ x_{t-i} \end{bmatrix} + \mathbf{e}_t$$

where \mathbf{c} is a vector of constants, \mathbf{M}_i are the $n \times n$ coefficient (or autoregressive) mixing matrices and \mathbf{e}_t is the unobservable white-noise. For the results shown below, $p = 2$. We then use the standard F-test to determine significance.

In Table 4, we show the causal relationships identified among the out-links of the article *Earthquake*. Recall that there was a major earthquake in Australia during our time-window. We wish to understand which of the out-links appeared to be sensitive to this large change in interest in Earthquake. We use a significance cutoff of 0.01 and test for Granger causality among the time-series with $s = 4$.

7 Conclusion

PageRank is one of the most widely used network centrality measures. Our dynamical system reformulation of PageRank permits us to incorporate time-dependent teleportation in a relatively seamless manner. Based on the results presented here, we believe this is an interesting variation on the PageRank model. For instance, we can analyze certain choices of oscillating teleportation functions (Lemma 6). Our empirical results show that the maximum change in the transient rank values identifies interesting sets of pages. Furthermore, this method is simple to implement in an online setting using either a forward Euler or Runge-Kutta integrator for the dynamical system. We hope that it might find a use in online monitoring systems.

One important direction for future work is to treat the inverse problem. That is, suppose that the observed page views reflect the behavior of these random surfers. For-

Table 4 – Example of Causality in Wikipedia. We only consider pages with a $pval < 0.01$ as statistically significant. The page with values “caused” by Earthquake represent ideas related to earthquakes. All pages below are significant with $pval < 0.01$.

Earthquake Granger causes	p-value
Seismic hazard	0.003535
Extensional tectonics	0.003033
Landslide dam	0.002406
Earthquake preparedness	0.001157
Richter magnitude scale	0.000584
Fault (geology)	0.000437
Aseismic creep	0.000419
Seismometer	0.000284
Epicenter	0.000020
Seismology	0.000001

mally, suppose that we equate page views with samples of $\mathbf{x}(t)$. Then, the goal would be to find $\mathbf{v}(t)$ that produces this $\mathbf{x}(t)$. This may not be a problem for websites such as Wikipedia, due to our argument that the majority of page views reflect search engine traffic. But for many other cases, we suspect that $\mathbf{x}(t)$ may be much easier to observe.

References

- S. Abiteboul, M. Preda, and G. Cobena. Adaptive on-line page importance computation. In *WWW*, pages 280–290. ACM, 2003. Cited on page 16.
- N.K. Ahmed, A.F. Atiya, N. El Gayar, and H. El-Shishiny. An empirical comparison of machine learning models for time series forecasting. *Econ. Rev.*, 29(5-6):594–621, 2010. Cited on page 22.
- Reid Andersen, Fan Chung, and Kevin Lang. Local graph partitioning using PageRank vectors. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science*, 2006. URL <http://www.math.ucsd.edu/~fan/wp/localpartition.pdf>. Cited on page 2.
- J.P. Bagrow, D. Wang, and A.L. Barabási. Collective response of human populations to large-scale emergencies. *PloS one*, 6(3):e17680, 2011. Cited on page 16.
- B. Bahmani, R. Kumar, M. Mahdian, and E. Upfal. PageRank on an evolving graph. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 24–32. ACM, 2012. Cited on page 15.
- Luca Becchetti, Carlos Castillo, Debora Donato, Ricardo Baeza-Yates, and Stefano

- Leonardi. Link analysis for web spam detection. *ACM Trans. Web*, 2(1):1–42, February 2008. ISSN 1559-1131. doi: 10.1145/1326561.1326563. Cited on page 10.
- Abraham Berman, Michael Neumann, and Ronald J. Stern. *Nonnegative Matrices in Dynamic Systems*. Wiley, 1989. Cited on page 6.
- M. Bianchini, M. Gori, and F. Scarselli. Inside PageRank. *ACM Transactions on Internet Technologies*, 5(1):92–128, 2005. ISSN 1533-5399. doi: 10.1145/1052934.1052938. Cited on page 15.
- Paolo Boldi. TotalRank: Ranking without damping. In *WWW*, pages 898–899, 2005. Cited on page 17.
- Paolo Boldi, Massimo Santini, and Sebastiano Vigna. Paradoxical effects in PageRank incremental computations. *Internet Mathematics*, 2(2):387–404, 2005. Cited on page 15.
- Paolo Boldi, Roberto Posenato, Massimo Santini, and Sebastiano Vigna. Traps and pitfalls of topic-biased PageRank. In *WAW2006, Fourth International Workshop on Algorithms and Models for the Web-Graph*, LNCS, pages 107–116. Springer-Verlag, 2007. doi: 10.1007/978-3-540-78808-9_10. Cited on page 2.
- G.E.P. Box, G.M. Jenkins, and G.C. Reinsel. *Time series analysis: forecasting and control*, volume 734. Wiley, 2011. Cited on page 25.
- S. Chien, C. Dwork, R. Kumar, D.R. Simon, and D. Sivakumar. Link evolution: Analysis and algorithms. *Internet Mathematics*, 1(3):277–304, 2004. Cited on page 16.
- Paul G. Constantine and David F. Gleich. Random alpha PageRank. *Internet Mathematics*, 6(2):189–236, September 2010. doi: 10.1080/15427951.2009.10129185. Cited on pages 4, 9, 10, and 12.
- A. Das Sarma, S. Gollapudi, and R. Panigrahy. Estimating PageRank on graph streams. In *SIGMOD*, pages 69–78. ACM, 2008. Cited on page 15.
- Daniel M. Dunlavy, Tamara G. Kolda, and Evrim Acar. Temporal link prediction using matrix and tensor factorizations. *TKDD*, 5(2):10:1–10:27, February 2011. ISSN 1556-4681. doi: 10.1145/1921632.1921636. Cited on page 16.
- Mark Embree and Richard B. Lehoucq. Dynamical systems and non-hermitian iterative eigensolvers. *SIAM Journal on Numerical Analysis*, 47(2):1445–1473, 2009. doi: 10.1137/07070187X. Cited on page 15.
- D. Gleich, P. Glynn, G. Golub, and C. Greif. Three results on the PageRank vector: eigenstructure, sensitivity, and the derivative. *Web Information Retrieval and Linear Algebra Algorithms*, 2007. Cited on page 17.

- D.F. Gleich, P.G. Constantine, A.D. Flaxman, and A. Gunawardana. Tracking the random surfer: empirically measured teleportation parameters in PageRank. In *WWW*, pages 381–390. ACM, 2010. Cited on pages 12 and 17.
- C.W.J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pages 424–438, 1969. Cited on page 25.
- P. Grindrod, M.C. Parsons, D.J. Higham, and E. Estrada. Communicability across evolving networks. *Physical Review E*, 83(4):046120, 2011. Cited on pages 4 and 16.
- Zoltán Gyöngyi, Hector Garcia-Molina, and Jan Pedersen. Combating web spam with TrustRank. In *Proceedings of the 30th International Very Large Database Conference*, Toronto, Canada, 2004. ISBN 0-12-088469-0. URL <http://i.stanford.edu/~zoltan/publications/gyongyi2004combating.pdf>. Cited on page 2.
- Roger A. Horn and Stefano Serra-Capizzano. A general setting for the parametric Google matrix. *Internet Mathematics*, 3(4):385–411, March 2007. URL <http://projecteuclid.org/euclid.im/1227025007>. Cited on pages 4 and 9.
- W. Kutta. Beitrag zur näherungsweise integration totaler differentialgleichungen. 1901. Cited on page 14.
- Amy N. Langville and Carl D. Meyer. Updating PageRank with iterative aggregation. In *WWW*, pages 392–393, 2004. Cited on page 16.
- Amy N. Langville and Carl D. Meyer. *Google’s PageRank and Beyond: The Science of Search Engine Rankings*. Princeton University Press, 2006. ISBN 978-0-691-12202-1. Cited on pages 1, 2, and 4.
- J.P. LeSage. Applied econometrics using MATLAB. *Manuscript, Dept. of Economics, University of Toronto*, 1999. Cited on page 25.
- J. O’Madadhain and P. Smyth. EventRank: A framework for ranking time-varying networks. In *LinkKDD*, pages 9–16. ACM, 2005. Cited on pages 10 and 16.
- L. Page, S. Brin, R. Motwani, and T. Winograd. The PageRank citation ranking: Bringing order to the web. 1998. Cited on pages 1 and 3.
- J. Ratkiewicz, S. Fortunato, A. Flammini, F. Menczer, and A. Vespignani. Characterizing and modeling the dynamics of online popularity. *PRL*, 105(15):158701, 2010. Cited on page 16.
- Ryan A. Rossi and David F. Gleich. Dynamic pagerank using evolving teleportation. In Anthony Bonato and Jeannette Janssen, editors, *Algorithms and Models for the Web*

- Graph*, volume 7323 of *Lecture Notes in Computer Science*, pages 126–137. Springer Berlin Heidelberg, 2012. ISBN 978-3-642-30540-5. doi: 10.1007/978-3-642-30541-2_10. Cited on pages 15 and 21.
- C. Runge. Über die numerische auflösung von differentialgleichungen. *Mathematische Annalen*, 46(2):167–178, 1895. Cited on page 14.
- Lawrence F. Shampine and Mark W. Reichelt. The MATLAB ODE suite. *SIAM Journal on Scientific Computing*, 18(1):1–22, 1997. doi: 10.1137/S1064827594276424. Cited on page 14.
- Rohit Singh, Jinbo Xu, and Bonnie Berger. Pairwise global alignment of protein interaction networks by matching neighborhood topology. In *Proceedings of the 11th Annual International Conference on Research in Computational Molecular Biology (RECOMB)*, volume 4453 of *Lecture Notes in Computer Science*, pages 16–31, Oakland, CA, 2007. Springer Berlin / Heidelberg. doi: 10.1007/978-3-540-71681-5_2. Cited on page 2.
- Jimeng Sun, Dacheng Tao, and Christos Faloutsos. Beyond streams and graphs: dynamic tensor analysis. In *SIGKDD, KDD '06*, pages 374–383, New York, NY, USA, 2006. ACM. ISBN 1-59593-339-5. doi: 10.1145/1150402.1150445. Cited on page 16.
- Hanghang Tong, Christos Faloutsos, and Jia-Yu Pan. Fast random walk with restart and its applications. In *ICDM '06: Proceedings of the Sixth International Conference on Data Mining*, pages 613–622, Washington, DC, USA, 2006. IEEE Computer Society. ISBN 0-7695-2701-9. doi: 10.1109/ICDM.2006.70. Cited on page 2.
- Panayiotis Tsaparas. Using non-linear dynamical systems for web searching and ranking. In *Proceedings of the twenty-third ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems, PODS '04*, pages 59–70, New York, NY, USA, 2004. ACM. ISBN 158113858X. doi: 10.1145/1055558.1055569. Cited on page 15.
- Various. Wikipedia database dump, 2009. Version from 2009-03-06. http://en.wikipedia.org/wiki/Wikipedia:Database_download. Cited on page 17.
- Various. Wikipedia pageviews, 2011. Accessed in 2011. <http://dumps.wikimedia.org/other/pagecounts-raw/>. Cited on page 17.
- Rebecca S. Wills and Ilse C. F. Ipsen. Ordinal ranking for Google’s PageRank. *SIAM Journal on Matrix Analysis and Applications*, 30:1677–1696, January 2009. doi: 10.1137/070698129. Cited on page 15.
- Jaewon Yang and Jure Leskovec. Patterns of temporal variation in online media. In *Proceedings of the fourth ACM international conference on Web search and data mining*,

WSDM '11, pages 177–186, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0493-1. doi: 10.1145/1935826.1935863. Cited on page [24](#).