Variable Damping Effect on Network Propagation

by

Yuchen Qian

Department of Computer Science
Duke University

Date: _____

Approved:

_____
Xiaobai Sun, Supervisor

_____
Nikos P. Pitsianis

_____
Jeffrey S. Chase

Thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science in the Department of Computer Science
in the Graduate School of Duke University
2018

Abstract

Variable Damping Effect on Network Propagation

by

Yuchen Qian

Department of Computer Science
Duke University

Date: _____

Approved:

_____
Xiaobai Sun, Supervisor

_____
Nikos P. Pitsianis

_____
Jeffrey S. Chase

An abstract of a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Department of Computer Science in the Graduate School of Duke University
2018

Abstract

In modern network analysis, the PageRank algorithm has been used as an indispensable tool to determine the importance and relevance of the network nodes. Inspite of extensive research conducted to accelerate the algorithm or its variants, there are few studies about the effects of the damping factors on the ranking distribution. To understand how the damping factor can affect the rank distribution in different PageRank models, specifically, the directed surfer model by Brin and Page, and the heat-kernel PageRank by Chung. We studied the ranking vector (steady state distribution) under different damping factor values with each model. Enabled by efficient batch calculation of the ranking vectors, we conducted systematic experiments to measure the discrepancies of the distributions, explored and explained the capability of adjusting the steady-state distribution via the change in damping factors. Experimental results show that the steady-state distribution by Brin-Page model responses non-linearly to the change in damping factor α, while by Chung's heat-kernel model, the damping factor β casts negligble effect on steady-state distribution. With this phenomenon, Brin-Page model may be preferable over Chung's model on utilizing the non-linear relationship between the damping factor and steady-state distribution. The relationship can be utilized also to find the propagation speed(damping factor) from observations of two or more consecutive distributions.

Key words: PageRank, Damping factor, KL divergence, Distribution

Contents

[1] X. Chen. Exploiting common structures across multiple network propagation schemes. 2017

List of Tables

List of Figures

List of Abbreviations and Symbols

Symbols

$a, b, c, \cdots$	Scalars
$A, M, P, \cdots$	Matrices
$\|\|A\|\|_F$	Frobenius norm of a matrix
A^{-1}	Inverse of a matrix
A^T	Transpose of a matrix
$f(A)$	A function apply on a matrix
$\rho(A)$	Spectral radius of a matrix
$\kappa(A, b)$	Krylov subspace with matrix A and vector b
A_{ij}	The i^{th} row and j^{th} column of a matrix.
G	Graph
(V, E)	Graph with V as vertices and E as edges.
x^*	Optimal solution for a linear system
e	Error vector

Abbreviations

dmperm	DulmageMendelsohn decomposition

Acknowledgements

I want to first express my great appreciation to my advisor Prof. Xiaobai Sun for her excellent advice and academic guidence. She has been consistently helping me in every aspect since the beginning of thesis including the theoretical analysis and experiment conducting. More importantly, while working with her, I learned how to conduct research, doing literature review, and other knowledge which I cannot learn from taking courses. I feel very lucky that I have her as my advisor.

I also want to share my thanks to my committee members Prof. Nikos P. Pitsianis and Prof. Jeffery S. Chase. I really appreciate their agreeing to serving on my committee, and they are the one of the reasons that I keep improving the standard of my thesis.

In addition to Professors, I would like to thank Xi Chen being my partner as we work together with Prof. Xiaobai Sun on different perspectives of the project. Xi has been always a wonderful colleague to work with as he provides proper advice and really contributes himself into the research process. As a friend, I would express my sincere thank to him.

I would like to thank my other twoo friends Tiancheng Liu and Zuoming Dai for any help during the research process. Tiancheng has been always giving me advice for research as well as any clarifications on the theory part while Zuoming helped me with his strong mathematical background.

Last but not least, I would greatly appreciate my parents for supporting me during

the completing of thesis. They gave me sincere advice and always encouraged me. Without them, I may not be able to continue my work.

1

Introduction

In early 1990s, people used websites like Yahoo, Lycos or AltaVista for search. These websites were all algorithmic search engines. In 1996, Larry Page and Sergey Brin developed an algorithm called PageRank, and 3 years later they published a paper describing the algorithm[19]. PageRank brought a novel concept of the importance of a webpage, and increased the accuracy of searching via the more accurate ordering based on the steady-state distribution of the random walk. Along with the invention of PageRank, Google was created, and it has become the most powerful web search engine in the world since then.

1.1 Brin-Page model and Personalized PageRank

Brin and Page[19], in their original paper, proposed the random surfer model and their PageRank algorithm. In the model, they assumed people surf randomly on the web following the hyperlinks on the webpage, and with a probability in certain form to jump to other nodes, not necessarily following the page links. The damping factor was introduced in this on-link and off-link random walk probability. Details

of this model and damping factor will be discussed in later sections. Following the initial work of Brin and Page[19], many academic papers have been published on PageRank or web search engine, and multiple aspects have been studied. (i) Accelerating PageRank computation. In the paper "Extrapolation methods for accelerating PageRank"[14], kamvar et al. used Aitken and quadratic extrapolation techniques. They reduced the time for reaching a residual of 0.001 by 23%. (ii) Increasing query accuracy. "Topic-sensitive PageRank"[12] by Haveliwala used "biased" PageRank vectors for calculation which increased the accuracy of querying. (iii) Customizing PageRank distribution, i.e., personalized PageRank. The paper "Scaling Personalized Web Search"[13] transformed personalized view of importance as partial vectors which could be shared across multiple personalized PageRank problems. The dynamic programming technique proposed in the paper improves the efficiency and accuracy of PageRank.

An important aspect in Personalized PageRank is the damping factor, which is based on the random surfer model by Brin and Page. In previous work[1], Boldi et al. analyzed the results of PageRank when damping factor is equal 0 and 1. The authors also proposed the idea that distribution could be considered as a function of damping factor. In [2], Bressan and Peserico discussed the relationship between damping factor and top k page ranking. Given the volume of work people have done for the above perspectives, the damping factor α, although has been studied in some publications[1, 2], needs to be studied more throughfully and interpreted more insightfully. In this thesis, we mainly investigated the effect of damping models on PageRank ranking distribution and its behavior, which hasn't been analyzed before.

1.2 Chung's model

An alternative model, heat-kernel PageRank, has been proposed and analyzed for around 15 years. Back in 2002, diffusion kernel on graphs was analyzed[15]. In 2007, Chung[6] introduced the PageRank based on discrete heat kernel. The parameter t in the heat kernel indicates the temperature. For heat-kernel PageRank, she introduced the paramter β as the diffusion speed for the network propagation. She used this PageRank model for solving local graph partition problem[7] and local linear systems with boundary conditions[8].

1.3 Intellectual merits and contribution

In this thesis, we report our investigation on the effect of damping models and parameters on ranking distributions. We consider two specific damping models, the Brin-Page model and Chung's model. With experiments, we will show that with Brin-Page model, a small change of damping factor α will make a great difference in ranking distribution. The damping factor β in the Chung's model does not contribute to personalizing – the distribution changes very little.

Chapter 2 describes the Brin-Page PageRank model, the personalized PageRank damping factor, and the Chung's heat-kernal PageRank model. Chapter 3 provides the network datasets we used throughout our experiment process. Chapter 4 introduces different approaches for solving the PageRank problem, and introduced a new method for efficient batch calculation of the ranking vectors from a brother thesis.[1] Chapter 5 analyzes the effect of damping factor with respect to different PageRank models. Chapter 6 concludes.

[1] X. Chen. Exploiting common structures across multiple network propagation schemes. 2017

2

Problem Description

2.1 PageRank

Assume we have a network graph $G = (V, E)$ with vertex set V and edge set E. Each node v_i represents a webpage, and an edge $e = (v_1, v_2)$ means that page v_1 has a hyperlink pointing to page v_2. What we want is to find out which set of nodes are "important". The importance of each node is described by its rank $R(v)$.

In their paper "The PageRank Citation Ranking: Bringing Order to the Web"[19] introduced a mode for the page ranks: The rank of page v is related to the ranks of its neighbor pages on the incoming links,

$$R(v) = \sum_{u \in P_v} \frac{R(u)}{N_u} \tag{2.1}$$

where P_v denotes the set of pages that point to page v, and N_u is the number of outgoing links from page u.

To be more specific, Equation (2.1) means that a rank of a page is the average rank of all pages that points to that specific page. Intuitively, this means that if a page is pointed by higher rank pages, it could have a high rank also. This is the basic idea

of PageRank.

2.2 Random Walk and Markov Chain

In this section, we will review PageRank in another perspective. PageRank can be actually considered as the random walk process on a Markov chain. A Markov chain is basically a stochastic model which describes a sequence of possible events. It is associated with a transitional matrix P, which is stochastic, i.e., its row sum or column sum is equal to 1. In this thesis, we define our P as column stochastic, i.e., $\sum_i P_{ij} = 1$ for every j. With an arbitrary initial ranking distribution x_0, we can always reach to a final steady ranking distribution x_f, assuming the network is connected and irreducible. Solving the PageRank problem could be further considered as finding a steady state probability distribution $x_f = Px_f$ of the Markov chain by random walk as follows:

$$x^{k+1} = Px^k \tag{2.2}$$

where x^k is the ranking distribution vector at k^{th} step in the random walk. We consider the steady state distribution is reached when x^{k+1} and x^k is close enough. (i.e. $||x^{k+1} - x^k||_\infty \leqslant \tau$). This walking precedure can be considered as the propagation of the initial configuration via the network.

2.3 Personalized PageRank

Brin and Page modified the random surfer model in [3]. Instead of following the network connection every time, each node can have an aggregated probability of $1-\alpha$ to jump to user-prefered nodes in the graph. They introduced a set of "trusted" nodes $\{v_i\}$. At each step, the destination of random walk is based on the distribution of v, and they gave them the name as source nodes. With these concepts, one can

construct a new Markov chain model for PageRank:

$$M = \alpha P + (1 - \alpha) v e^T \tag{2.3}$$

where e^T is a constant-1 vector. The new Markov model assigns each node an aggregate probability of α following existing hyperlinks(graph edges) and an aggregate probability of $1 - \alpha$ to choose a random node in the "trusted" set v according to user defined distribution. The parameter α is named as the "Damping Factor" with range in (0, 1]. In reality, a person would not always click the links in a page. He or she may like to jump into any website as he or she prefers. The damping factor in the model mimics this reality. It also "damps" or reduces the network propagation impact with the number of walk steps. A smaller value of α indicates that the network propagation damps faster. This model is also named *directed surfer model* by Richardson and Domingos[20].

For the same network, the desired damping speed may be different depending on the application. The "trusted" nodes may also be different from one application to another. Personalized PageRank is thus introduced. Personalized Pagerank, as its name suggested, customized the network propagation behavior by setting up specific distribution vector v and damping factor α.

2.4 Chung's PageRank model

An alternative PageRank model, the heat-kernel PageRank, was introduced by Fan Chung [6]. Here, heat-kernel is related a diffusion process. In general, heat will transfer from high temperature to low temperature, and if restricted within a fixed region, eventually the temperature will be uniformly distributed. A heat kernel is simply a solution to a specific diffusion problem.

Heat kernel is associated with the heat equation

$$\frac{\partial}{\partial t}x(\beta, v) = -x(\beta, v)(I - P) \tag{2.4}$$

The solution to the heat equation provides the heat-kernel PageRank:

$$x = e^{-\beta L}v = e^{-\beta}\sum_{i}^{\infty}\frac{(\beta)^k}{k!}(P^k v) \tag{2.5}$$

where $L = I - P$. We denote $H_t = e^{-\beta L}$ as the heat kernel to the graph, where β determines the diffusion speed. In the following contents, we will use β as the time parameter for heat-kernel PageRank. More details about the exponential sum form can be found in the Appendix A.6.

Notice that these two different PageRank models have the same intuition: each finds a final steady state distribution, with respect to a damping factor. They differ in the manner how the random walk vector $P^k v$ are summed together: one with geometric damping weights α^k, and the other with exponentially damping weights $\frac{\beta^k}{k!}$. In Chapter 5, we will provide an experiment process for studying these two different PageRank models, and evaluating how damping factors could impact on the final distribution.

3

Network Datasets and characteristics

3.1 Data Statistics

Table 3.1 provides a summary of the 6 network datasets which we used in our study. These datasets all come from KONECT[17]. The largest strongly connected components is denoted as LSCC. Our-degree of a node means how many nodes it points to. Diameter means by a certain number of steps, a node can reach to all of the other nodes in the corresponding network. All of the networks have a relative small diameter.

Table 3.1: Dataset Characteristics

	Total #nodes	#nodes in LSCC	out-degree	diameter
Google[1]	875,713	434,818	8.86	28
Wikipedia[2]	12,150,976	7,283,915	50.48	8
dbpedia[3]	18,268,992	3,796,073	26.76	8
Twitter(www)[4]	41,652,230	33,479,734	42.65	9
Twitter(mpi)[5]	52,579,682	40,012,384	47.57	12
friendster[6]	68,349,466	48,928,140	32.76	22

3.2 Preprocessing: System reduction

The original dataset contains two columns: FromID and ToID, which represent in row i the source and destination nodes of the i^{th} edge. We performed Dulmage-Mendelsohn(DM) decomposition[10] of the graph. This decomposition renders graph into strongly connected components. In this paper, we use the largest strongly connected component of the graph to conduct analysis of PageRank computation and Personalized PageRank; dangling nodes are eliminated and not considered.
Figure 2.1 visualized the network data after the node permutation by the Dulmage-Mendelsohn decomposition. The largest diagonal (square) block corresponds to the largest strongly connected component. The components of the graph are weakly connected. As we can see from these graphs, all the networks have a large connected component.

[1] A network of web pages connected by hyperlinks, 2002

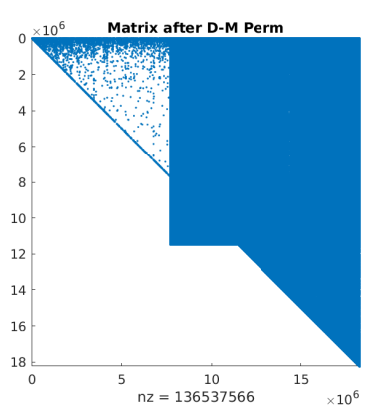
[2] Wikilinks of the Wikipedia in the English language

[3] Hyperlink network of Wikipedia extracted in DBpedia

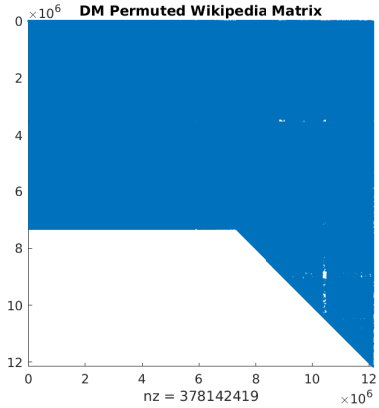
[4] Twitter follower network, 2010

[5] Twitter snapshot follower network, 2009

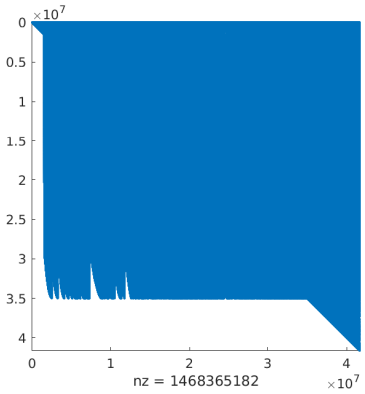
[6] Friendship network from Friendster(a social gaming cite), 2011



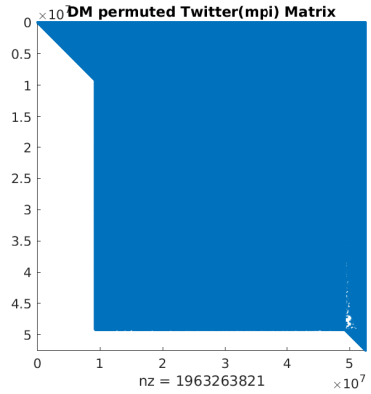
(a) DBpedia



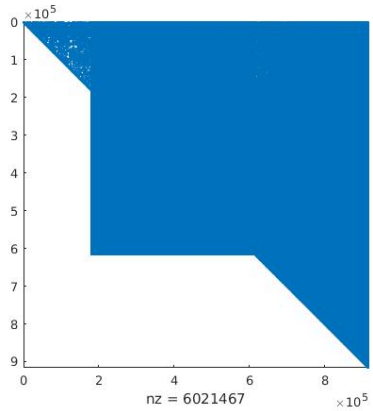
(b) Wikipedia



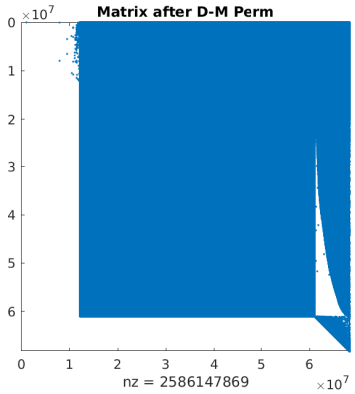
(c) Twitter-www



(d) Twitter-mpi



(e) Google



(f) Friendster

FIGURE 3.1: Network components revealed by Dulmage-Mendelsohn decomposition

4

Network Propagation Analysis

In this chapter, we reformulate the network propagation problem of the Brin-Page damping model introduced in Chapter 2, as (i) Perron-Frobenius eigenvalue problem associated with the personalized Markov chain (random walk graph); (ii) Equivalent linear system; and (iii) Power series expansion. It is worth noting that the solution to Chung's heat kernel damping model [6] takes the form of power series expansion as well. We therefore integrates the heat kernel damping model problem formulation into the corresponding subsection.

Following the reformulation of the damping models, we studied different approaches for obtaining the steady-state distributions and utilizing different problem formations.

Moreover, to enable an efficient and acccurate study of the effects of different damping models, and the effects of variant damping parameters in each of the models, we introduce the Krylov method to obtain the steady-state distribution (ranking vector x) in a batch-wise manner.

4.1 Damping model reformulation

4.1.1 Perron-Forbenius eigenvalue problem

PageRank problem, in specific, directed surfer model [20], is equivalent to seeking the steady-state solution to the Markov chain

$$M = \alpha P + (1 - \alpha)ve^T \tag{4.1}$$

where v is the personalized distribution, and $e^T x = 1$. The steady-state solution x satisfies the following equilibrium:

$$M(\alpha, v)x = x \tag{4.2}$$

As per Perrson-Frobenius theorem, the steady-state solution is the eigenvector corresponding to the largest eigenvalue $\lambda_{max} = 1$.

4.1.2 Equivalent linear system

By Equations (4.1) and (4.2), we have

$$x = \alpha Px + (1 - \alpha)ve^T x \tag{4.3}$$

$$(I - \alpha P)x = (1 - \alpha)v \tag{4.4}$$

which is the equivalent linear system form of the equilibrium (4.2). This reformulation enables the utilization of conventional method for solving linear system.

4.1.3 The steady-state distribution expressed by power series

Based on Equation (4.4), we have

$$x = (I - \alpha P)^{-1}(1 - \alpha)v \tag{4.5}$$

One way to expand $(I - \alpha P)^{-1}$ is to use Neumann series[18]

$$(I - \alpha P)^{-1} = \sum_{k=0}^{\infty} (\alpha P)^k \tag{4.6}$$

which transforms the Brin-Page model into a geometric sum[1] (power series) of random walks.

According to Chung[6], heat-kernal PageRank could be also represented in this format:

$$x = e^{-\beta L} v_0 = e^{-\beta(I-P)} v_0 = e^{-\beta} e^{\beta P} v_0 \tag{4.7}$$

and we can expand $e^{\beta P}$ by:

$$e^{\beta P} = \sum_{k=0}^{\infty} \frac{(\beta P)^k}{k!} \tag{4.8}$$

which takes the form of exponential sum[2].

4.2 Solving for the steady-state distribution

4.2.1 Power method for Perron-Forbenius eigenvalue problem

A solution to eigenvalue problem is by random walk. Given a Markov chain matrix $M(\alpha, v)$, we can compute the distribution at k^{th} step by the following:

$$x_{k+1} = M x_k$$

Therefore, we keep multiplying this matrix $M(\alpha, v)$(keep walking) until we are very close to the equilibrium (steady-state). To be more precise, we terminate our calculation when the norm of difference of two consecutive ranking vectors reaches certain tolerance level. This is the power method[11], which is mainly used for obtaining the eigenvector corresponding to the largest eigenvalue 1.

4.2.2 Iterative solutions to linear system

The reformulated linear system (4.3) allows us to view the problem from a different perspective. Recall we have the following equation:

$$(I - \alpha P)x = b \tag{4.9}$$

[1] See Appendix A.5 for more details

[2] See Appendix A.6 for more details

where $b = (1 - \alpha)v$. To solve the linear system (4.9), which is large and sparse, we can utilize conventional iterative methods such as Jacobi method and Gauss-Seidel method. In the Jacobi method[9], we split a matrix into two parts: $A = D - N$. D is the diagonal matrix, and N is the off-diagonal matrix. In our case, $D = I$ and $N = \alpha P$. We then rewrite the fixed point equilibrium and turn it into iteration as follows[3]:

$$Dx_{k+1} = Nx_k + b$$

In order for the Jacobi method to converge, the spectral radius of error propagation matrix must less than 1, i.e.

$$\rho(D^{-1}N) < 1$$

The proof of convergence is noted in the Appendix[4]. With the PageRank system, Jacobi method is guaranteed to converge.

The Gauss-Seidel method[9] can be also applied to this linear system. The idea behind Gauss-Seidel method is roughly the same as the Jacobi method as they both split the original matrix into several parts. For Gauss-seidel Method, we split a matrix into three parts: $A = D - L - U$. D is the diagonal matrix, L is the lower triangular matrix, and U is the upper triangular matrix. We then reformulate the equilibrium and turn it into iteration as follows[5]:

$$(D - L)x_{k+1} = Ux_k + b$$

Similar to the convergence analysis of Jacobi method, the spectral radius of error propagation matrix must less than 1, i.e.

$$\rho((D - L)^{-1}U) < 1$$

[3] See Appendix A.2 for detailed expansion

[4] See Appendix A.2

[5] See appendix A.3

The prove of convergence is also in the Appendix.[6]

4.2.3 Binomial expansion[7]

At k iteration, we have:

$$x_k = M^k x_0 \tag{4.10}$$

By the binomial expansion with Equation (4.10):

$$x_k = \sum_{i=0}^{k-1} \binom{k}{i} \alpha^i (1-\alpha)^{k-i} P^i v + \alpha^k P^k x_0 \tag{4.11}$$

From Equation(4.11) we could see that a change of α would only lead to a change in the binomial coefficients of each term $P^i v$. Therefore, storing $P^i v$ for each term would help avoiding matrix-vector re-calculation. We try to find a sufficiently large k with different damping factors, which guarantees the accuracy of the approximation. For Google network graph matrix, that k is chosen to be 500, according to the experimental results under different α values (0.5 to 1 with an interval of 0.05). However, this method doesn't yield correct results in practice due to rounding errors. For most of the terms of the binomial expansion, $\binom{k}{i}$ is an overly large integer that can not be precisely stored as a 8-byte unsigned integers while $\alpha^i(1-\alpha)^{k-i}$ is an overly small number that can not be precisely stored as a 8-byte double. Rounding errors are accumulating during multiplication and addition, and eventually cause incorrect results.

4.3 Krylov space method

We utilize Krylov space method for solving the steady-state solutions:

$$x_\alpha = (1-\alpha)(I - \alpha P)^{-1} v_0 \tag{4.12}$$

[6] See Appendix A.3

[7] X. Chen. Exploiting common structures across multiple network propagation schemes. 2017

$$x_\beta = e^{-\beta} e^{\beta P} v_0 \tag{4.13}$$

The theoretical and technical details can be found in a brother thesis from Xi Chen[4]. Briefly speaking, Krylov method relies on he pre-calculated invariant subspace $\kappa_m(P, v)$, allowing fast, batch-wise calculation under different damping models and damping factors. Via the fast calculation of functions on the compact representation (Hessenberg matrix H), we can calculate the linear systme withouu directly involving the original large matrix P. We will show in later sections the cost comparison (courtesy of Xi Chen) between Gauss-Seidel method and Krylov space method demonstrating the efficiency of Krylov space method.

4.4 Experimental comparison in computation cost

Table 4.1 shows the comparison result between Krylov method and Gauss-seidel method. We use matrix-vector multiplication counts as comparison criteria. For

Table 4.1: Comparison in matrix-vector production counts

	Krylov method	Gauss-Seidel
DBpedia	600	66
Wikipedia	600	102
Twitter(www)	600	72
Twitter(mpi)	600	106
Friendster	600	82

Krylov method, the multiplication count is mainly for generating the Krylov subspace. On the other hand, the number of counts for Gauss-seidel stands for the multiplication count of PageRank calculation with a single damping factor. Thus, if we want to calculate PageRank with respect to n α values, for DBpedia the count will be $66n$. It is not hard to interpret that if $n >= 10$, the Krylov method will out-perform Gauss-Seidel method in terms of efficiency(running-time). For other

datasets, n could be even smaller. Therefore, if we want to calculate Personalized PageRank with multiple damping factors, Krylov method will be more efficient.

5

Effects of different damping models and factors

The effects of damping models and their parameter values on the steady-state distribution rankings shall be better undetstood. In this paper, we conduct experiments on two particular models: Brin-Page model and Chung's model, and analyze their behaviours and the response in ranking to variation in the parameter of each model.

5.1 Evaluation Criteria

5.1.1 Ranking as Distribution

The ranking vector v is actually a probability distribution over the page nodes. Each value v_i in the distribution indicates the certain probability of the corresponding node being visited. In order to mearure the discrepancy in the steady-state distributions with different damping factor values with each model, we use the Kullback-Leiber Divergence as an aggregrated and stochastic evaluation measurement, rather than the top-k nodes ranking as in [2], which is sensitive to noise due to the nature of combinatorial comparison at micro-scale level.

5.1.2 Kullback-Leiber Divergence

Kullback-Leiber Divergence (KL-score) was introduced in 1959 by Kullback and Leiber[16]. It describes the "Information Discrimination" of two distributions. To be more specific, for two continuous distributions P and Q, the mean information discrimination from Q to P is defined as follows:

$$KL(P||Q) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)} d\lambda(x) \tag{5.1}$$

The discrete counterpart is:

$$KL(P||Q) = \sum_{i}^{n} p(x) \log \frac{p(x)}{q(x)} \tag{5.2}$$

We utilize the KL-score (5.2) to measure the discrepancies between the steady-state distributions corresponding to different damping models and factor values.

5.1.3 Rate of Change

We introduce a differential measurement D_{KL} for damping factor analysis – the rate of change in KL-score to the change of damping factor. We denote v_α to the ranking distribution with damping factor value α. If we change α to $\alpha + \Delta\alpha$, we will have another distribution $v_{\alpha+\Delta\alpha}$. The rate of change is:

$$D_{KL}(v_\alpha||v_{\alpha+\Delta\alpha}) = \frac{KL(v_\alpha||v_{\alpha+\Delta\alpha})}{\Delta\alpha} \tag{5.3}$$

When $\Delta\alpha$ is small, $D_{KL}(v_\alpha||v_{\alpha+\Delta\alpha})$ is an approximation of $\frac{dKL}{d\alpha}$, providing the distribution sensitivity measure with respect to α. In other words, $D_{KL}(v_\alpha||v_{\alpha+\Delta\alpha})$ provides an intuitive way to probe the effect of the damping factor within each model, which in turn, characterizes the tunability of the model.

5.2 Experiment Results

We describe in Table (5.1) our experiment setting:

Table 5.1: Experimental settings for comparison between different PageRank model

	α	$\Delta\alpha$	reference α	reference α
Brin-Page	$0.7 \sim 1$	0.001, 0.005, 0.01, 0.02	0.85	1
	β	$\Delta\beta$	reference β	reference β
Chung	$0.7 \sim 1$	0.001, 0.005, 0.01, 0.02	0.85	1

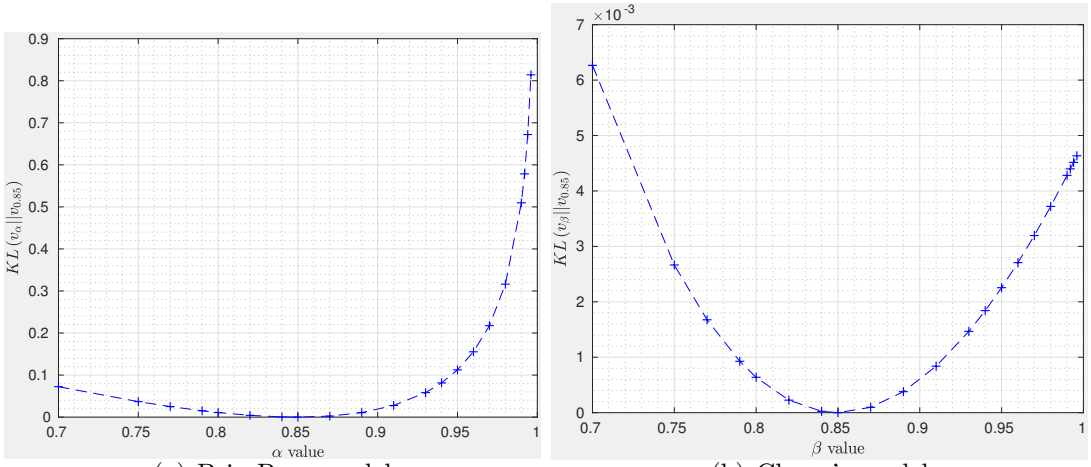
We set the model parameters α and β to be in the same range, by the 1st-order terms in respective Taylor expansions,

$$(I - \alpha P)^{-1} = \sum_{k=0}^{\infty} (\alpha P)^k = I + (\alpha P)^1 + (\alpha P)^2 + \cdots + (\alpha P)^i + \cdots$$

$$e^{\beta P} = \sum_{k=0}^{\infty} \frac{(\beta P)^k}{k!} = I + (\beta P)^1 + \frac{(\beta P)^2}{2!} + \cdots + \frac{(\beta P)^i}{i!} + \cdots$$

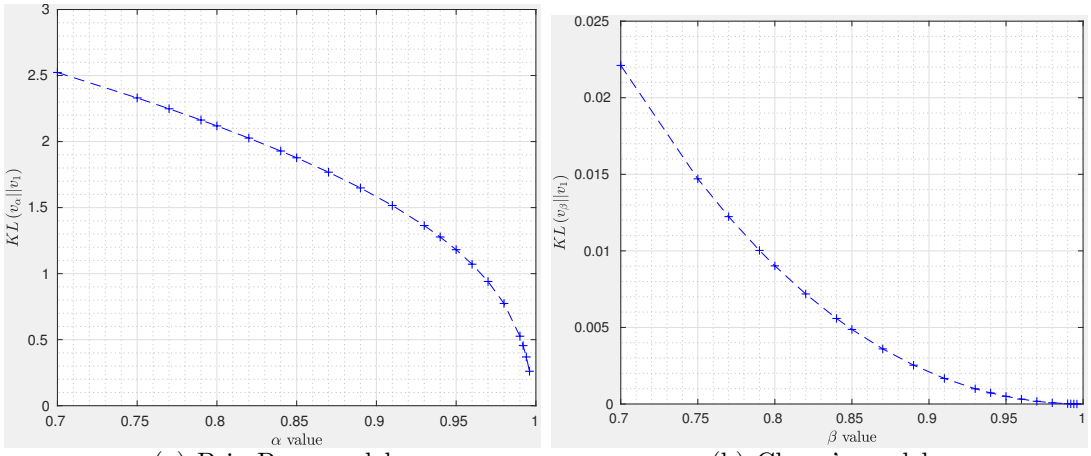
5.2.1 Comparison to Reference Ranking Distributions

We first looked at the relationship between the KL score and damping models. The results with Google dataset are shown in Figures 5.1 and 5.2.



(a) Brin-Page model (b) Chung's model

FIGURE 5.1: Comparison in KL-divergence scores between the cases of $\alpha, \beta = 0.85$ with Google network



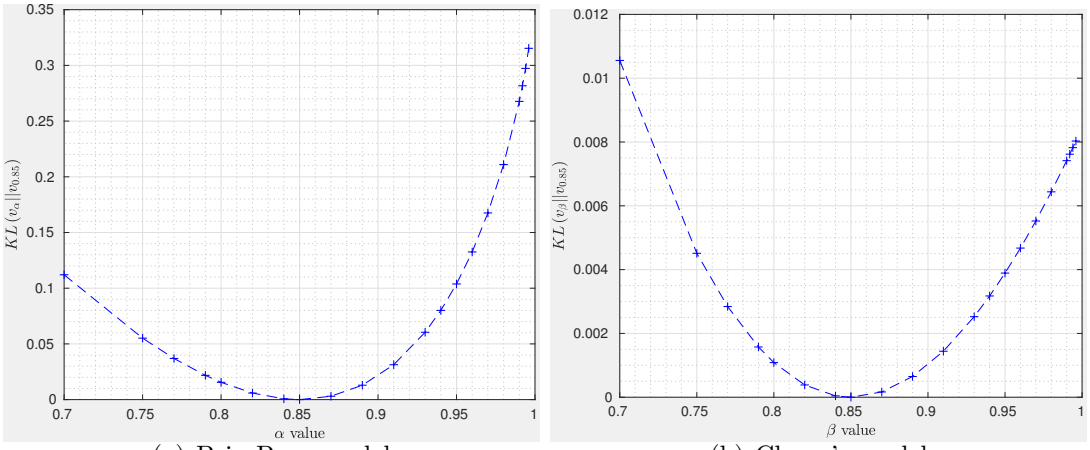
(a) Brin-Page model (b) Chung's model

FIGURE 5.2: Comparison in KL-divergence scores between the cases of $\alpha, \beta = 1$ with Google network

Figures 5.1 shows the discrepancy between a reference steady-state distribution at a fixed reference value 0.85 and ranking distributions at other damping factor values, in [0.7, 1]. In Figure 5.1, we choose $\alpha = 0.85$ to generate the reference ranking vector. The specific reference choice of α value here follows the proposal of Brin and Page[19]. In Figure 5.2, we choose $\alpha = \beta = 1$ as the reference value. The steady-

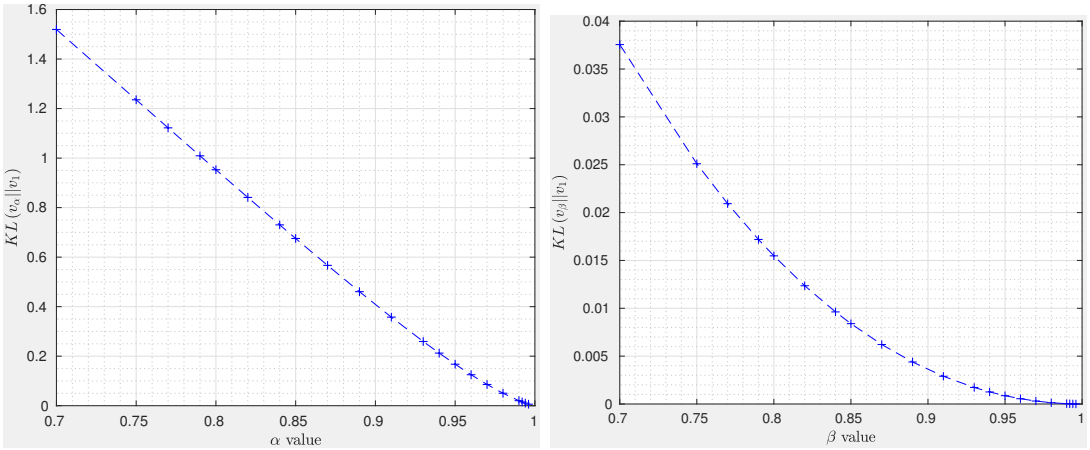
state distribution follows the non-personalized Markov model.

As shown in the both Figures, when the damping value is close to the reference value, the KL-score decreases, and as the value becomes far away from the reference value, the KL-score increases. By the measure definition, the KL score at 0.85 is zero. From Figure 5.1, we can see that the curve is not as symmetric as that in Figure 5.2. For Brin-Page model, as the relative value is close to 1 starting from 0.85, the KL score increases drastically. Notice that the magnitude of KL score for the two Figures are not the same. For Brin-Page model, the KL score ranges from (0.1, 0.9), while for Chung's model, the order of magnitude is at 10^{-3}.



(a) Brin-Page model (b) Chung's model

FIGURE 5.3: Comparison in KL-divergence scores between the cases of $\alpha, \beta = 0.85$ with Twitter-mpi network



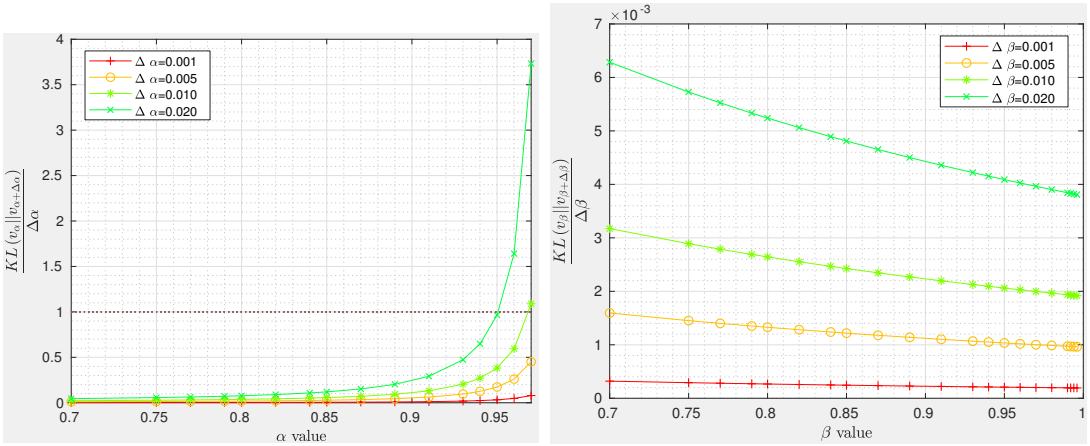
(a) Brin-Page model (b) Chung's model

FIGURE 5.4: Comparison in KL-divergence scores between the cases of $\alpha, \beta = 1$ with Twitter-mpi network

Figures 5.3 and 5.4 show the results with Twitter-mpi dataset. The results are very similar to the results of Google dataset. KL-score monotonic increases when damping factor becomes far away from the reference values.

5.2.2 Comparison in rate of change

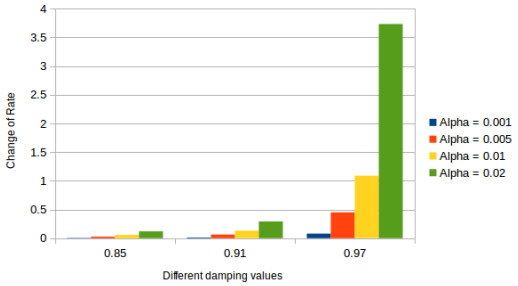
We analyzed the sensitivitiy to two damping models: Brin-Page and Chung's model. Results are shown in Figure 5.5.



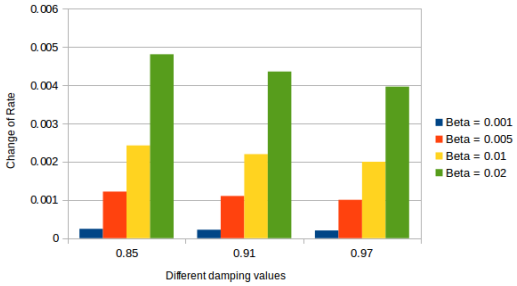
(a) Brin-Page model (b) Chung's model

FIGURE 5.5: Changing rate of KL-divergence in response to incremental change in damping model parameters with Google network

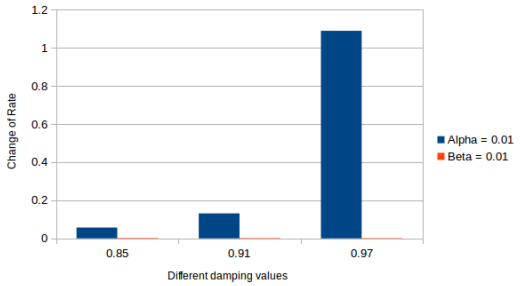
By Figure 5.5(a), D_{KL} increases as $\Delta\alpha$ increases. Also, D_{KL} changes dramatically when α is approaching to 1. In contrast to Figure 5.5(a), 5.5(b) shows that the change rate of KL-score decreases as $\Delta\beta$ increases. It is worth mentioning that the $D_{KL}(\beta)$ is significantly smaller than $D_{KL}(\alpha)$, which implies that a small change in the α will result in a great change in the ranking distribution and the opposite for β.



(a) $KL(v_\alpha||v_{\alpha+\Delta\alpha})$ with Brin-Page model



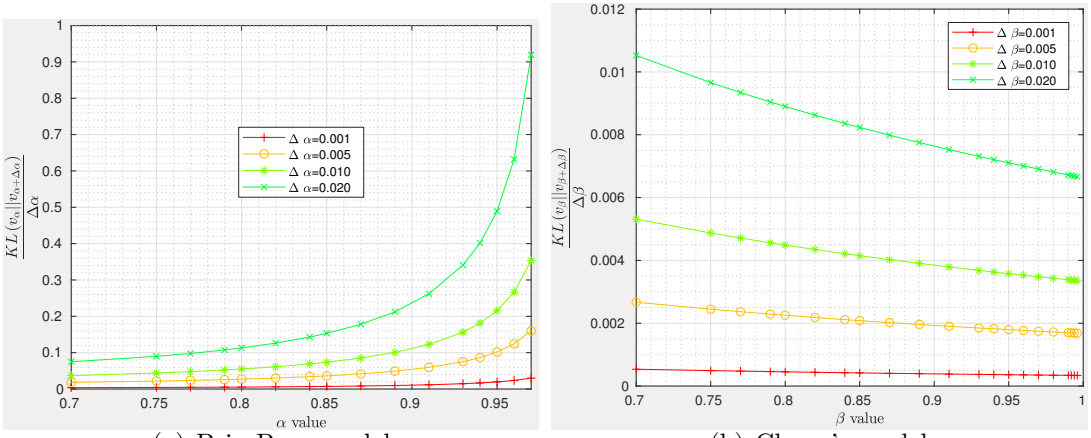
(b) $KL(v_\beta||v_{\beta+\Delta\beta})$ with Chung's model



(c) $KL(v_\alpha||v_{\alpha+\Delta\alpha})$ with Brin-Page model and $KL(v_\beta||v_{\beta+\Delta\beta})$ with Chung's model

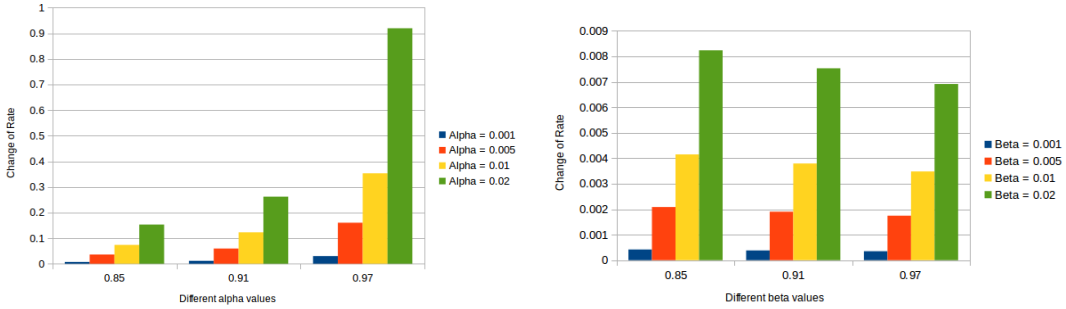
FIGURE 5.6: Changing rate of KL-divergence in response to incremental change in damping model parameters in histogram with Google network

To highlight the significant difference, Figure 5.6 provides another view of the sensitivity in histogram. Figure 5.6(c) presents the comparison between two damping models with $\Delta\alpha = \Delta\beta = 0.01$. Apparently, Chung's heat-kernel PageRank model is insensitive to the change of parameter β in the range [0.7, 1] comparing to Brin and Page model.

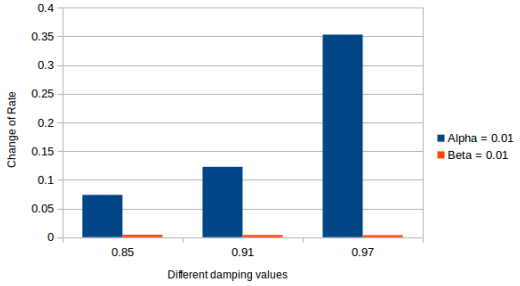


(a) Brin-Page model (b) Chung's model

FIGURE 5.7: Changing rate of KL-divergence in response to incremental change in damping model parameters with Twitter-mpi network



(a) $KL(v_\alpha||v_{\alpha+\Delta\alpha})$ with Brin-Page model (b) $KL(v_\beta||v_{\beta+\Delta\beta})$ with Chung's model



(c) $KL(v_\alpha||v_{\alpha+\Delta\alpha})$ with Brin-Page model and $KL(v_\beta||v_{\beta+\Delta\beta})$ with Chung's model

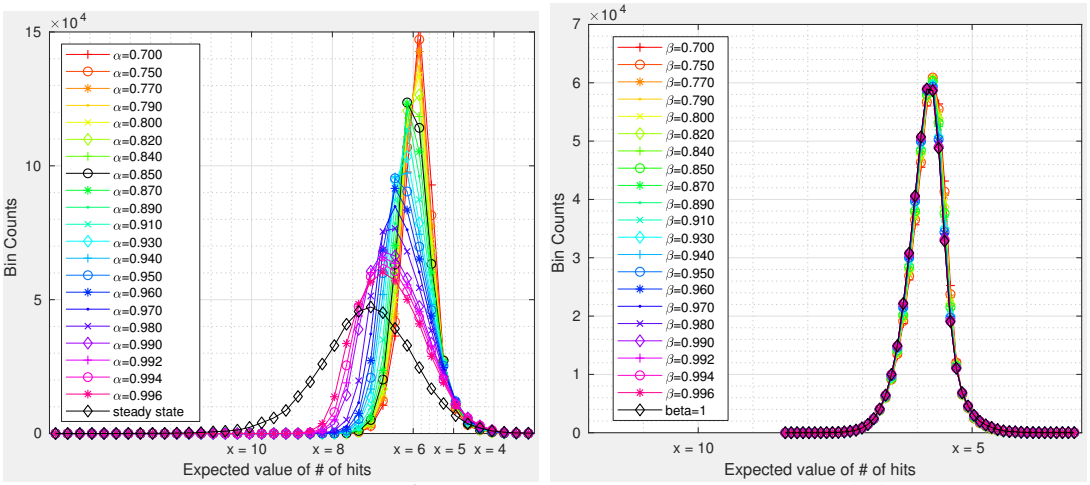
FIGURE 5.8: Changing rate of KL-divergence in response to incremental change in damping model parameters in histogram with Twitter-mpi network

Figure 5.7 and 5.8 shows the corresponding results with Twitter-mpi network.

The results are similar to that of Google network.

5.2.3 Absolute change in expected page visits

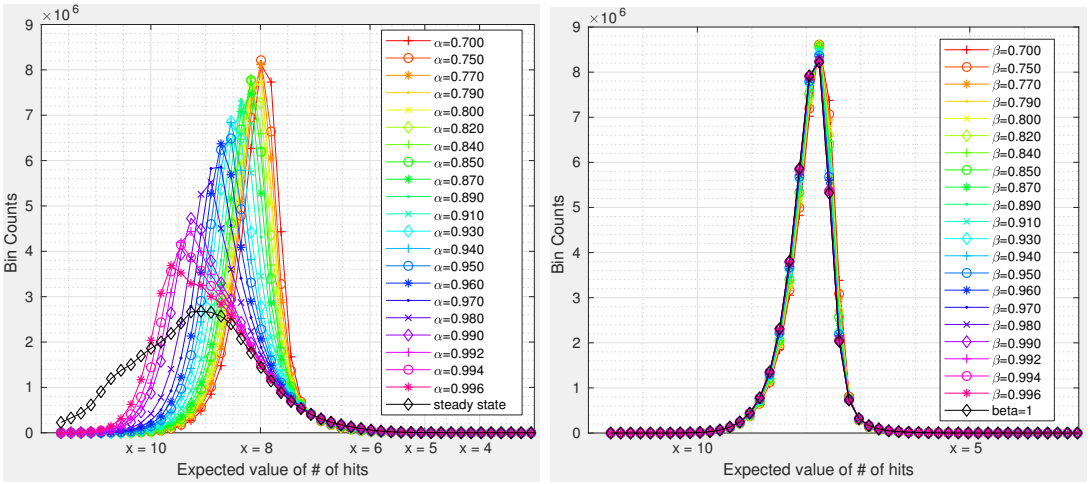
We present the distribution of expected page visits with respect to different damping models given the relationship between ranking and the expected page visits. In Figure 5.9, the x-axis denotes the normalized page visits. The y-axis indicates the corresponding number of nodes potentially from all other pagesin the same bin of expected visit number. The curve is in fact a histogram of expected clicks over the nodes.



(a) $v_\alpha = (1-\alpha)(I-\alpha P)^{-1}v_0$ (b) $v_\beta = e^{-\beta}e^{\beta P}v_0$

FIGURE 5.9: Changes in PageRank distribution with different damping parameter values, with Google network

The curves in Figure 5.9(a) are with respect to different α values with Brin-Page model. It is noticeable that different α value will result in different shapes of ranking distributions. When the α value is 0.7, it has the highest peak and narrowest support bandwidth. As α increases, the peak becomes lower and the distribution is flattening out. With Chung's model, the shape of distribution changes little with respect to β, see Figure 5.9(b). Such behaviors apply to the Twitter-mpi network, and the results are shown in Figure 5.10.

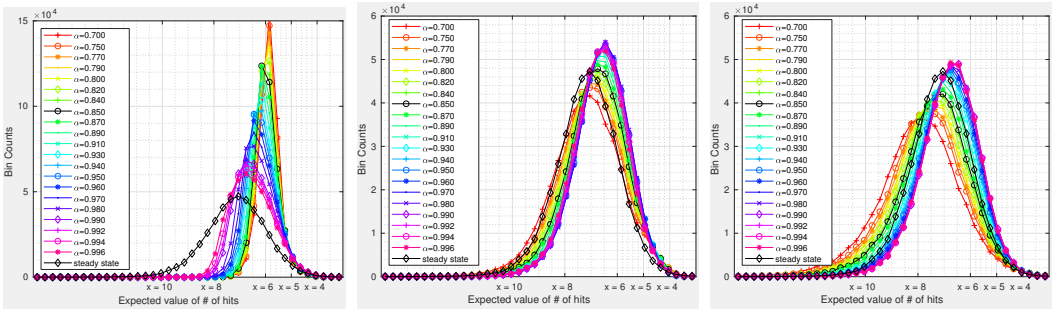


(a) $v_\alpha = (1-\alpha)(I-\alpha P)^{-1}v_0$ (b) $v_\beta = e^{-\beta}e^{\beta P}v_0$

FIGURE 5.10: Changes in PageRank distribution with different damping parameter values, with Twitter-mpi network

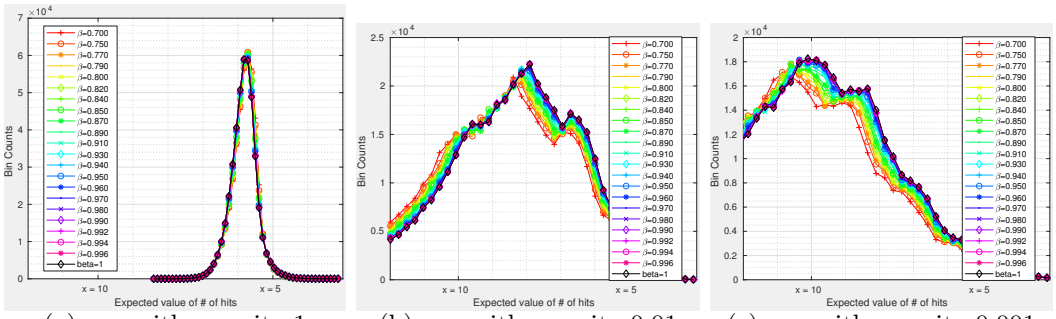
5.2.4 Response to sparsity in personalized vector

Both models depend also on the structure of the personalized vector v. We conducted the similar experiments to observe the response in page ranking to the change in the sparsity of the personalized vector using personalized vector v, i.e., the initial distribution vector, with different sparsity. Figure 5.11 and 5.12 shows the ranking distribution with respect to different initial vectors with different sparsity.



(a) $v_{\alpha 0}$ with sparsity 1 (b) $v_{\alpha 0}$ with sparsity 0.01 (c) $v_{\alpha 0}$ with sparsity 0.001

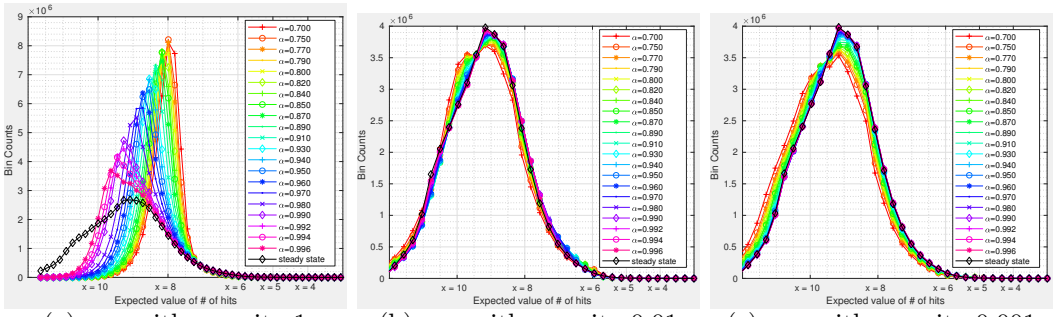
FIGURE 5.11: Changes in PageRank distribution with different sparsity of personalized distribution, with Google network



(a) $v_{\beta 0}$ with sparsity 1 (b) $v_{\beta 0}$ with sparsity 0.01 (c) $v_{\beta 0}$ with sparsity 0.001

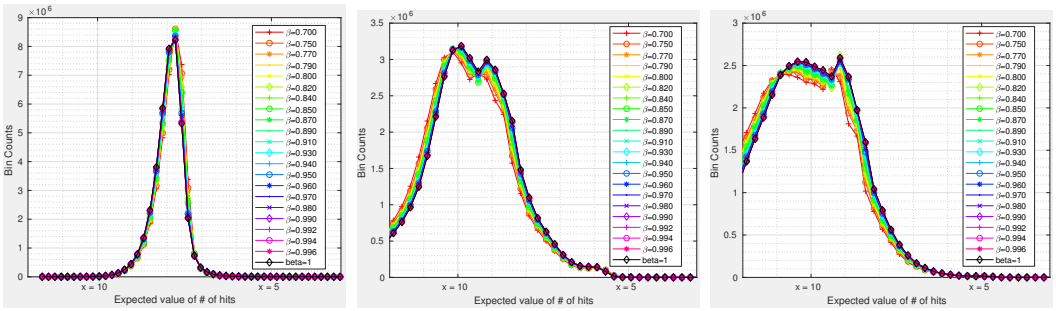
FIGURE 5.12: Changes in PageRank distribution with different sparsity of personalized distribution, with Google network

From Figure 5.11, we can see that as the sparsity for initial vector increase, the descrepancies between distributions under different damping factor α values suppresses (from 5.11(a) to 5.11(c)). This behavior applies to both damping models.



(a) $v_{\alpha 0}$ with sparsity 1 (b) $v_{\alpha 0}$ with sparsity 0.01 (c) $v_{\alpha 0}$ with sparsity 0.001

FIGURE 5.13: Changes in PageRank distribution with different sparsity of personalized distribution, with Twitter-mpi network



(a) $v_{\beta 0}$ with sparsity 1 (b) $v_{\beta 0}$ with sparsity 0.01 (c) $v_{\beta 0}$ with sparsity 0.001

FIGURE 5.14: Changes in PageRank distribution with different sparsity of personalized distribution, with Twitter-mpi network

The corresponding results of Twitter-mpi further indicates the effect of sparsity on the ranking distribution. If we increase the sparsity of the personalized distribution, we will obtain a smoother and flatter final distribution.

5.3 Summary of Experimental findings

We summarize our findings as follows:

1. The steady state distribution by Brin-Page model is sensitive to its associated damping factor α, while in Chung's heat-kernel model, the damping factor β casts negligble effect on steady-state distribution, in the range $\beta \in (0.7, 1]$.

2. In Brin-Page model, the damping factor α can change the normalized visiting number histogram substantially, while the damping factor β in Chung's model has little effect on the normalized visit numbers.

3. A sparser personalized vector v will result in less variation in steady-state distribution with varying damping parameter values.

According to the results shown above, by choosing different damping factors, we could customize the shape of the steady-state distribution. Also, given a desired disribution, we can change the damping factor to match the requirement.

6

Conclusion

In this thesis, we introduced different equilibrium forms of the PageRank system in detail. With respect to different forms, we provide corresponding solutions.

In addition, we conducted systematic experiments to measure the discrepancies of the distributions with respect to the change of damping factors in different damping models. The steady state distribution in Brin-Page model is sensitive to its associated damping factor α. On the other hand, for Chung's heat-kernel PageRank, the damping factor β does little effect on the steady-state distribution.

Last but not least, when the personalized vector v for Brin-Page model becomes sparser, we will have less variation in steady-state distribution of different damping factors. For the Chung's model, further research and experiment need to be conducted.

Appendix A

A.1 The Power Method

The pseodocode of power method is as follows:

initialize x_0, M, τ, r;
while $r < \tau$ **do**
  $x_{k+1} = Mx_k$;
  $r = ||x_{k+1} - x_k||$;
end
return x_{k+1}

Algorithm 1: Pseudocode for the power method

The power method will eventually converge to the principle eigenvector of M[14] under the condition that the spectral radius of M is smaller or equal to 1. The convergence rate will be

$$\frac{|\lambda_2|}{|\lambda_1|}$$

where

$$1 = |\lambda_1| > |\lambda_2| \geqslant \cdots \geqslant |\lambda_n| > 0$$

A.2 Jacobi Method

For Jacobi Method, We decompose matrix A into diagonal matrix D and off-diagonal matrix N. Therefore, we can reconstruct our linear system as follows:

$$(I - \alpha P)x = b$$

$$(D - N)x = b$$

$$Dx - Nx = b$$

$$x = Nx + b$$

$$x_{k+1} = Nx_k + b$$

Therefore, we only need to compute $\alpha P x_k + b$ for every iteration, which is very similar to what we do in the power method.

The pseudocode is as follows:

initialize x_0,threshold;
decompose A into D and N;
while $r < \tau$ **do**
  $Dx_{k+1} = Nx_k + b$;
  $r = ||x_{k+1} - x_k||$;
end
return x_{k+1}

Algorithm 2: Jacobi method pseudocode

The error propagation for Jacobi Method is as follows:

$$Dx_{k+1} = Nx_k + b \tag{A.1}$$

$$Dx^* = Nx^* + b \tag{A.2}$$

Subtract $A.2$ by $A.1$ we have

$$De_{k+1} = Ne_k \tag{A.3}$$

$$e_{k+1} = D^{-1}Ne_k \tag{A.4}$$

where e is the error between optimal solution and current solution. We can see that the propagation matrix is $D^{-1}N$, and if we could restrict the largest eigenvalue of the

matrix to be less that 1, the error will decrease every step, and Jacobi will eventually converge.

A.3 Gauss-seidel Method

For Gauss-seidel method, we split matrix $A = (I - \alpha P)$ into diagonal matrix D, lower off-diagonal matrix L and upper off-diagonal matrix U. Therefore, we can reconstruct our linear system with $\alpha P = L + U$ and $D = I$,

$$(I - \alpha P)x = b$$

$$(I - L - U)x = b$$

$$(I - L)x = Ux + b$$

$$(I - L)x_{k+1} = Ux_k + b$$

The pseudocode is as follows:

initialize x_0, τ r;
split A into D, L and U;
while $r < \tau$ **do**
    $(D - L)x_{k+1} = Ux_k + b$;
    $r_{k+1} = ||x_{k+1} - x_k||$;
end
return x_{k+1}

Algorithm 3: Pseudocode for the Gauss-seidel method

The error propagation is as follows:

$$(I - L)x_{k+1} = Ux_k + b \tag{A.5}$$

$$(I - L)x^* = Ux^* + b \tag{A.6}$$

Subtract $A.1$ by $A.2$ we have

$$(I - L)e_{k+1} = Ne_k \tag{A.7}$$

$$e_{k+1} = (I - L)^{-1} N e_k \tag{A.8}$$

where e is the error between optimal solution and current estimate. The propagation matrix is $(I - L)^{-1}U$. If the largest eigenvalue of the matrix is less that 1, the error will decrease every step, and the iteration converges.

A.4 Convergence Analysis of PageRank System

The PageRank system is

$$(I - \alpha P)x = b,$$

where $\alpha \in (0, 1)$, P is the probability matrix with zeros on the diagonal and for any column vector p_i of P, $||p_i||_1 = 1$. It is easy to verify that $I - \alpha P$ is a diagonally dominant matrix.

Before experiment, we need to show that the Guass-Seidel methods converges in solving our problem. To show the convergence, we first apply a lemma.

Lemma 1. *If matrix A is diagonally dominant, Q is a permutation matrix, then $Q^\top AQ$ is also diagonally dominant.*

Proof. Without loss of generality, we consider $Q = Q_{ij}$ where Q_{ij} is the result of swapping the ith and jth rows of I. If A can be written in a row, then

$$A = \begin{bmatrix} \dots \\ a_i \\ \dots \\ a_j \\ \dots \end{bmatrix}, Q^\top A = \begin{bmatrix} \dots \\ a_j \\ \dots \\ a_i \\ \dots \end{bmatrix}, Q^\top AQ = \begin{bmatrix} \dots \\ a_{j*} \\ \dots \\ a_i^* \\ \dots \end{bmatrix}$$

$Q^\top A$ does not change other rows and only swaps the ith and jth rows of A. $Q^\top AQ$ puts the a_{jj} to the ii position and puts a_{ii} to the jj position. Then the diagonal elements of A still stay as diagonal elements in $Q^\top AQ$; while the off-diagonal elements

stay with the same corresponding diagonal elements from original A. So, $Q^\top AQ$ keeps the diagonal dominance. □

Theorem 2. *If A is diagonally dominant, then, for system $Ax = b$, Guass-Seidel method converges.*

Proof. Consider decomposition $A = D - L - U$, where D, L, U are respectively diagonal, lower triangular and upper triangular matrices. For Guass-Seidel method, the iteration step is

$$(D - L)x_{k+1} = Ux_k + b.$$

Let $A(\lambda) = \lambda(D - L) - U$. If $|\lambda| \geqslant 1$, then $A(\lambda)$ is diagonally dominant, therefore, nonsingular. So, if

$$\det\Big(\lambda(D - L) - U\Big) = 0,$$

then $|\lambda| < 1$. So,

$$\rho((D - L)^{-1}U) < 1.$$

Therefore, Guass-Seidel method converges. □

These proofs can be found in [5]

A.5 PageRank as Geometric Sum

$$(I - \alpha P)^{-1} = \sum_{k=0}^{\infty} (\alpha P)^k$$

$$\begin{aligned}(I - \alpha P)\sum_{k=0}^{\infty}(\alpha P)^k &= ((\alpha P)^0 + (\alpha P)^1 + (\alpha P)^2 + \cdots + (\alpha P)^k) \\ &\quad - (\alpha P)^1 + (\alpha P)^2 + (\alpha P)^3 + \cdots + (\alpha P)^{k+1}) \\ &= (I - (\alpha P)^{k+1})\end{aligned}$$

$$\lim_{k\to\infty}(\alpha P)^{k+1} = Q(\alpha\Lambda)^{k+1}Q^T$$

$$(\alpha\Lambda)^{k+1} = \begin{bmatrix} (\alpha\lambda_1)^{k+1} & & \\ & \ddots & \\ & & (\alpha\lambda_n)^{k+1} \end{bmatrix}$$

$$where \lim_{k\to\infty}(\alpha\lambda_i)^{k+1} = 0$$

$$(I - (\alpha P)^{k+1}) = I$$

$$(I - \alpha P)\sum_{k=0}^{\infty}(\alpha P)^k = I$$

$$\sum_{k=0}^{\infty}(\alpha P)^k = (I - \alpha P)^{-1}$$

A.6 PageRank as Exponential Sum

What we have is:

$$e^x = \sum_{k=0}^{\infty}\frac{x^k}{k!}$$

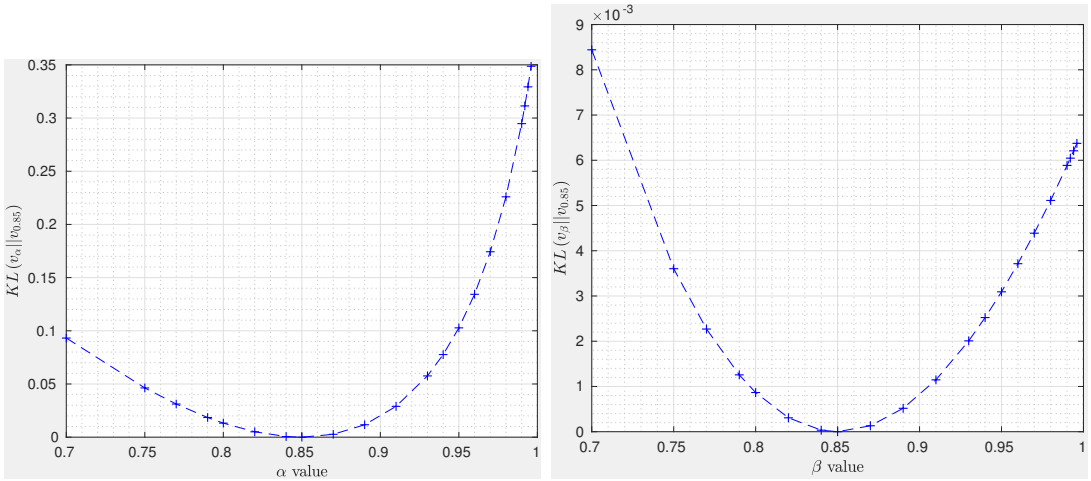
Need to show:

$$e^{\beta P} = \sum_{k=0}^{\infty}\frac{(\beta P)^k}{k!}$$

$$
\begin{aligned}
P &= Q\Lambda Q^T \\
e^{\beta P} &= Qe^{\beta\Lambda}Q^T \\
&= Q\begin{bmatrix} e^{\beta\lambda_1} & & \\ & \ddots & \\ & & e^{\beta\lambda_n} \end{bmatrix} Q^T = Q\begin{bmatrix} \sum_{k=0}^{\infty}\frac{(\beta\lambda_1)^k}{k!} & & \\ & \ddots & \\ & & \sum_{k=0}^{\infty}\frac{(\beta\lambda_n)^k}{k!} \end{bmatrix} Q^T \\
&= \sum_{k=0}^{\infty}\frac{(\beta)^k}{k!}Q\begin{bmatrix} (\lambda_1)^k & & \\ & \ddots & \\ & & (\lambda_1)^k \end{bmatrix} Q^T = \sum_{k=0}^{\infty}\frac{(\beta)^k}{k!}Q\Lambda^k Q^T \\
&= \sum_{k=0}^{\infty}\frac{(\beta P)^k}{k!}
\end{aligned}
$$

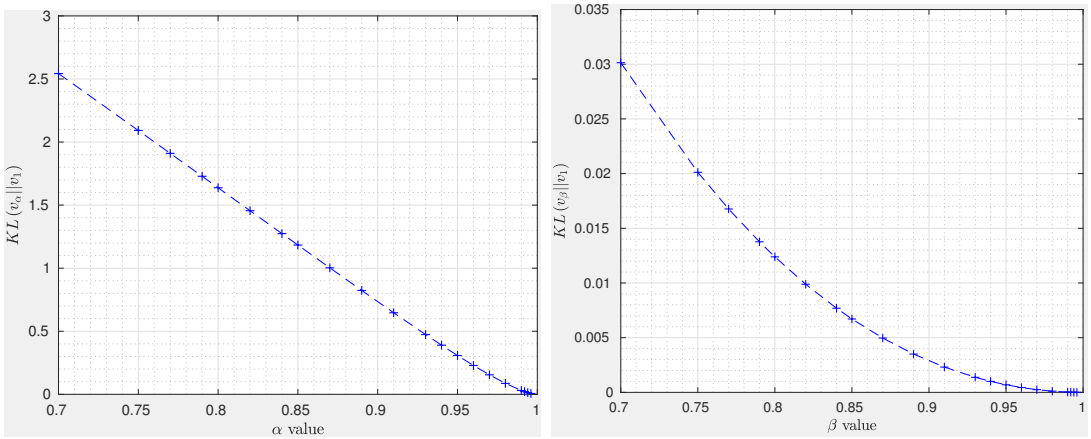
A.7 Supplementary results for Section 5.2

A.7.1 Wikipedia



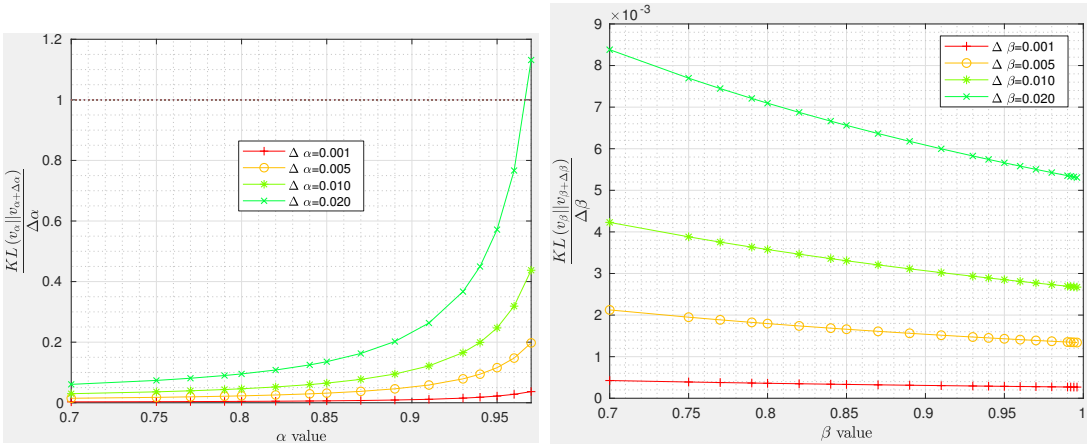
(a) Brin-Page model (b) Chung's model

FIGURE A.1: Comparison in KL-divergence scores between the cases of $\alpha, \beta = 0.85$ with Wikipedia network



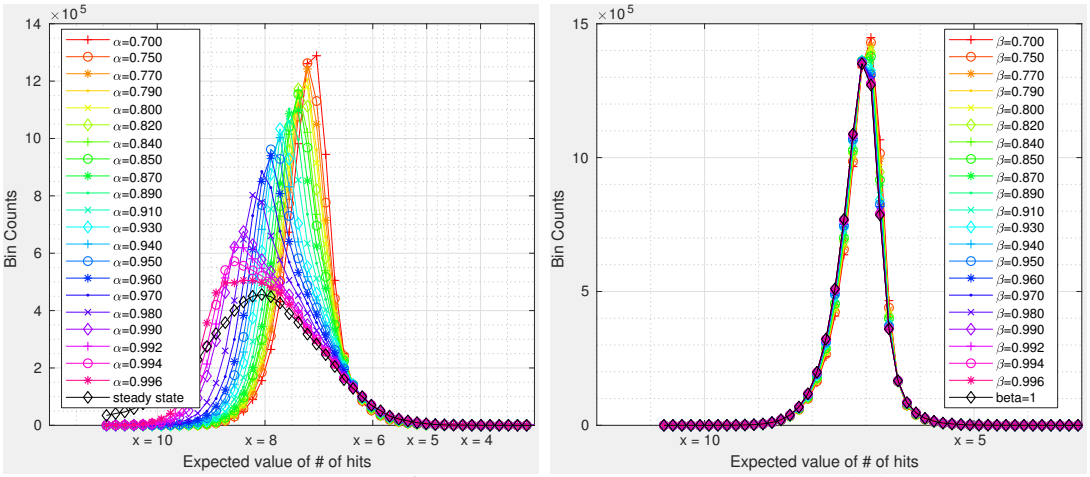
(a) Brin-Page model (b) Chung's model

FIGURE A.2: Comparison in KL-divergence scores between the cases of $\alpha, \beta = 1$ with Wikipedia network



(a) Brin-Page model (b) Chung's model

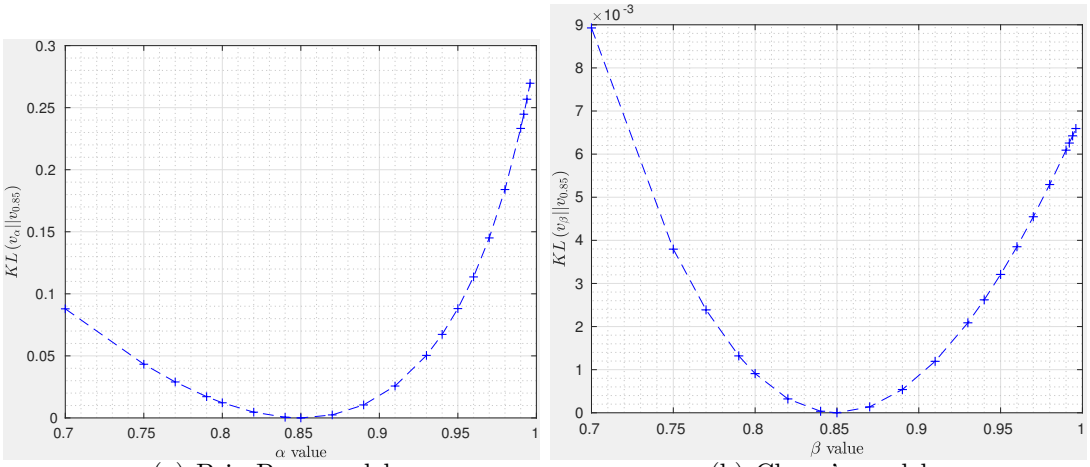
FIGURE A.3: Changing rate of KL-divergence in response to incremental change in damping model parameters with Wikipedia network



(a) $v_\alpha = (1-\alpha)(I-\alpha P)^{-1}v_0$ (b) $v_\beta = e^{-\beta}e^{\beta P}v_0$

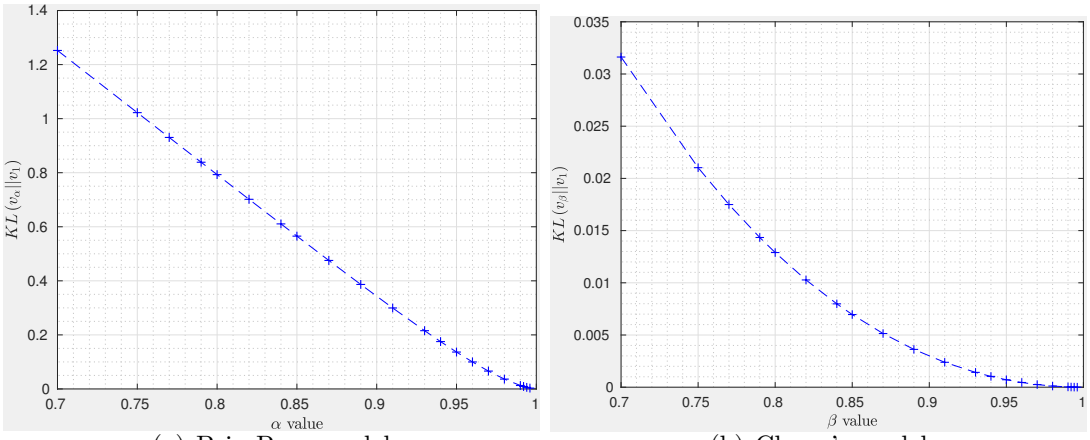
FIGURE A.4: Changes in PageRank distribution with different damping parameter values, with Wikipedia network

A.7.2 DBpedia



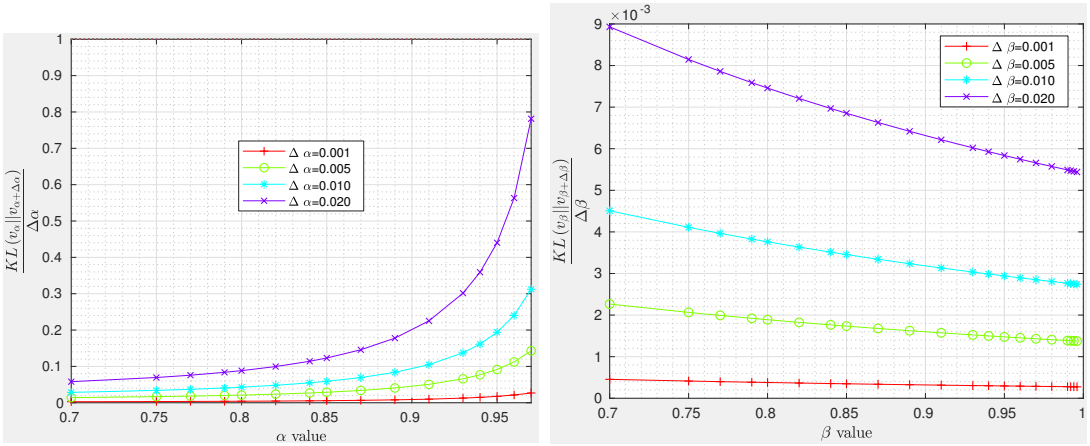
(a) Brin-Page model (b) Chung's model

FIGURE A.5: Comparison in KL-divergence scores between the cases of $\alpha, \beta = 0.85$ with DBpedia network



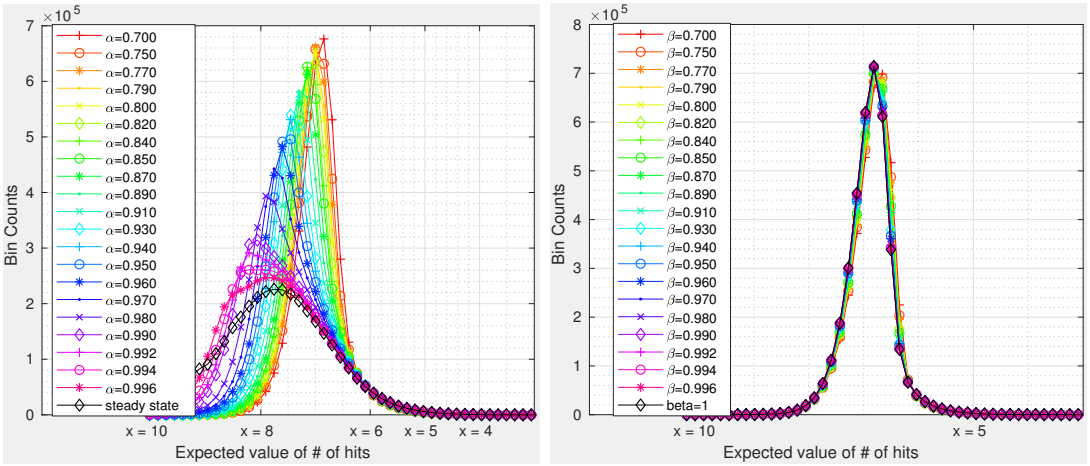
(a) Brin-Page model (b) Chung's model

FIGURE A.6: Comparison in KL-divergence scores between the cases of $\alpha, \beta = 1$ with DBpedia network



(a) Brin-Page model (b) Chung's model

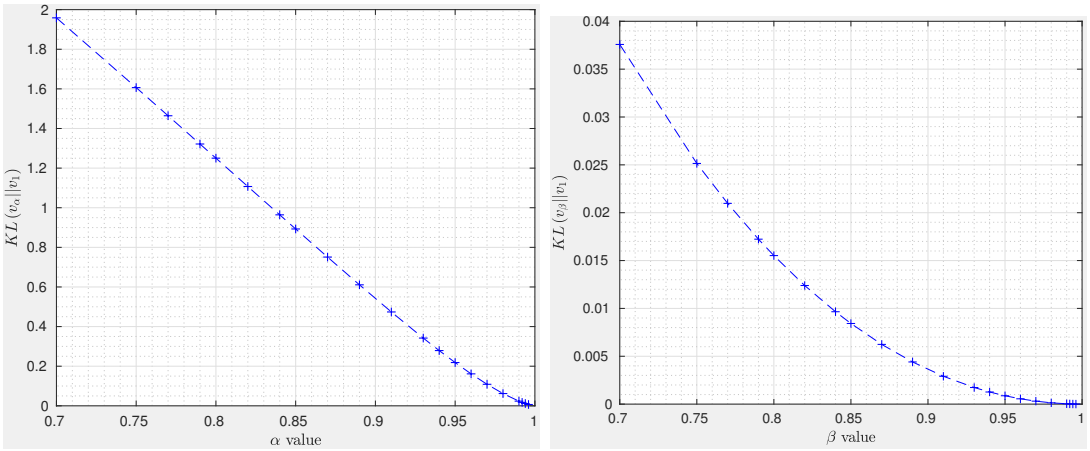
FIGURE A.7: Changing rate of KL-divergence in response to incremental change in damping model parameters with DBpedia network



(a) $v_\alpha = (1-\alpha)(I-\alpha P)^{-1}v_0$ (b) $v_\beta = e^{-\beta}e^{\beta P}v_0$

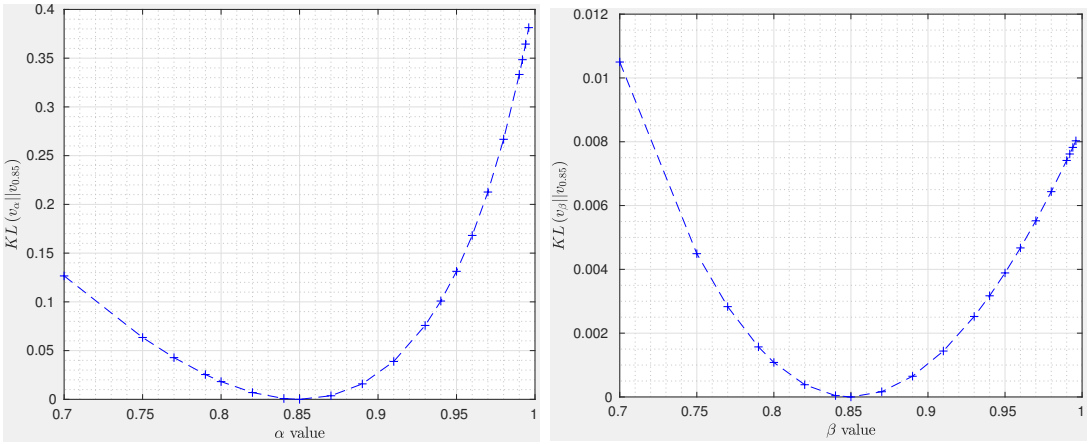
FIGURE A.8: Changes in PageRank distribution with different damping parameter values, with DBpedia network

A.7.3 Twitter-www



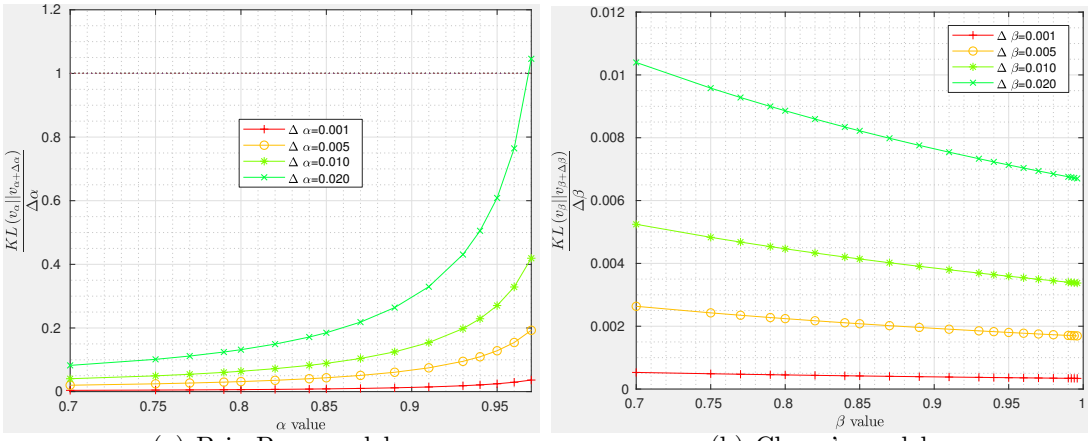
(a) Brin-Page model (b) Chung's model

FIGURE A.9: Comparison in KL-divergence scores between the cases of $\alpha, \beta = 0.85$ with Twitter-www network



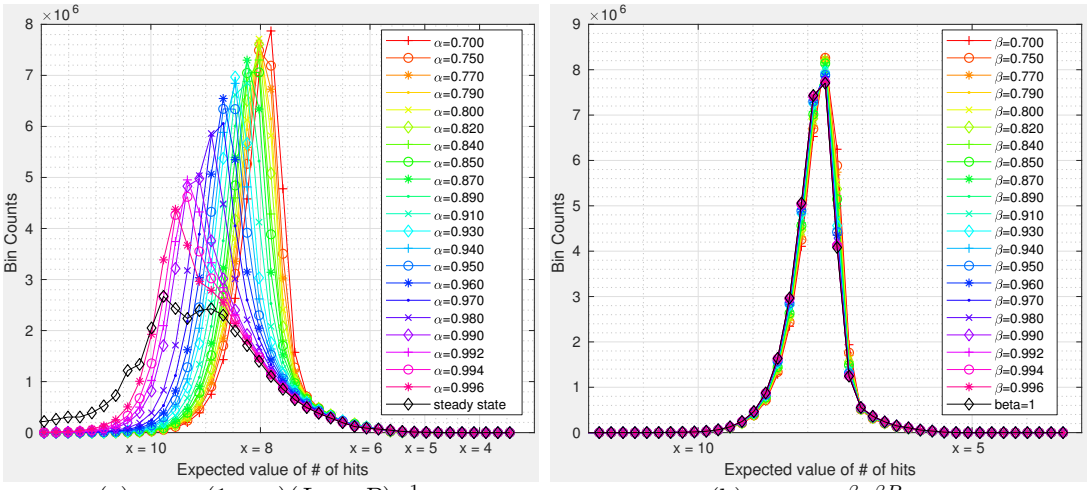
(a) Brin-Page model (b) Chung's model

FIGURE A.10: Comparison in KL-divergence scores between the cases of $\alpha, \beta = 1$ with Twitter-www network



(a) Brin-Page model (b) Chung's model

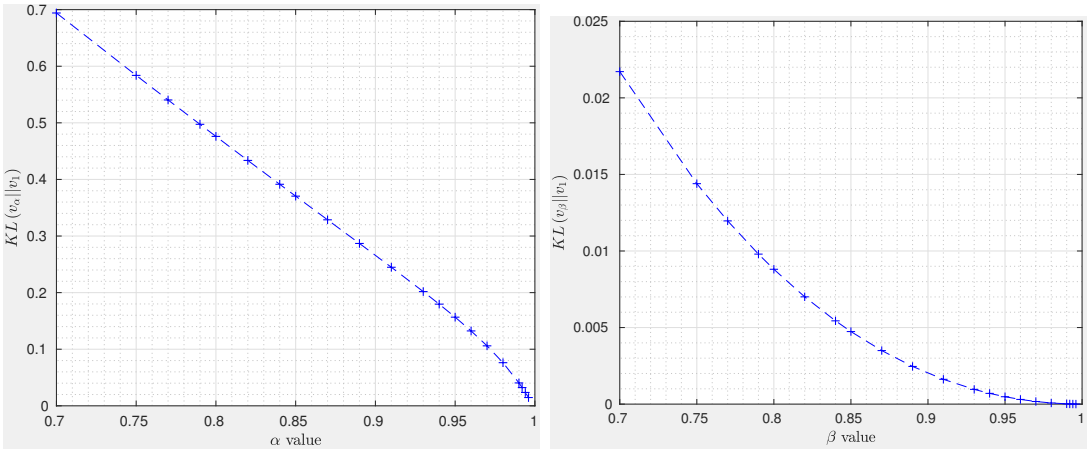
FIGURE A.11: Changing rate of KL-divergence in response to incremental change in damping model parameters with Twitter-www network



(a) $v_\alpha = (1-\alpha)(I-\alpha P)^{-1}v_0$ (b) $v_\beta = e^{-\beta}e^{\beta P}v_0$

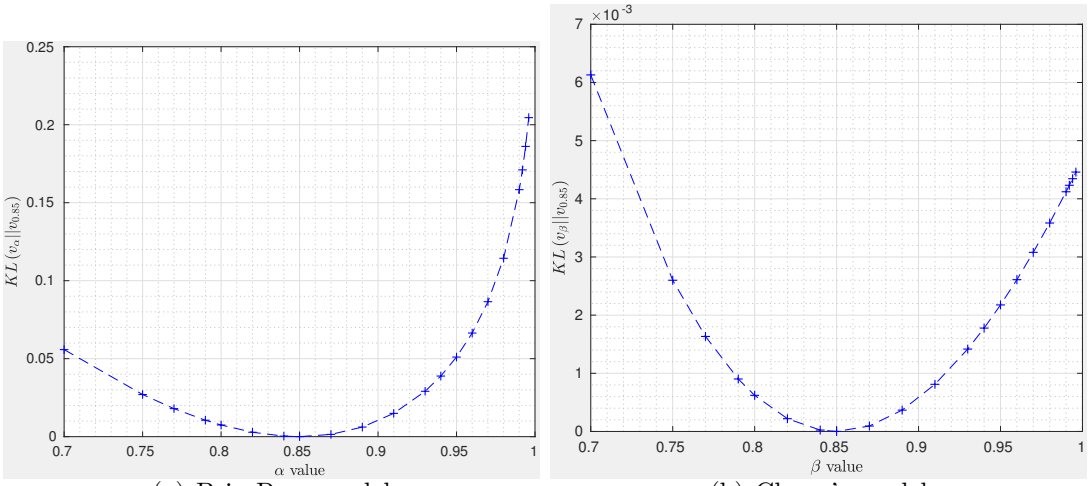
FIGURE A.12: Changes in PageRank distribution with different damping parameter values, with Twitter-www network

A.7.4 Friendster



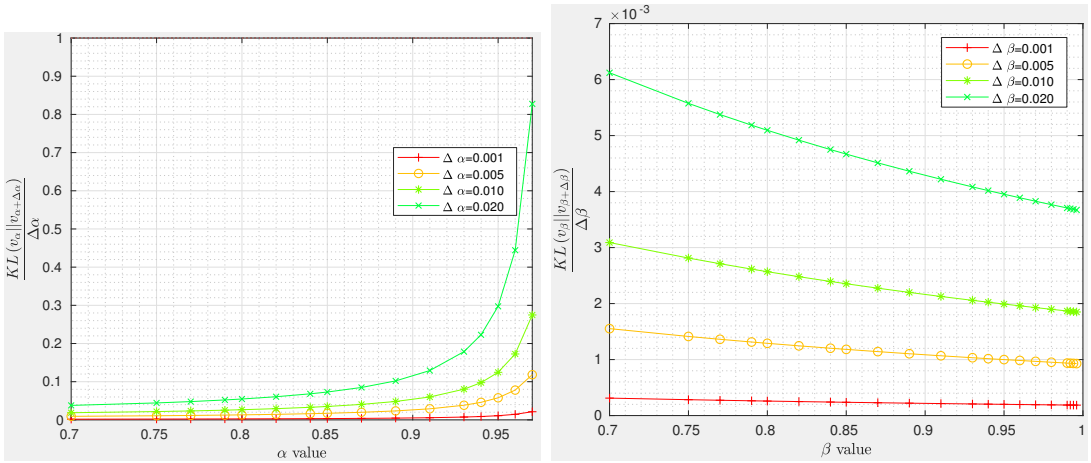
(a) Brin-Page model (b) Chung's model

FIGURE A.13: Comparison in KL-divergence scores between the cases of $\alpha, \beta = 0.85$ with friendster network



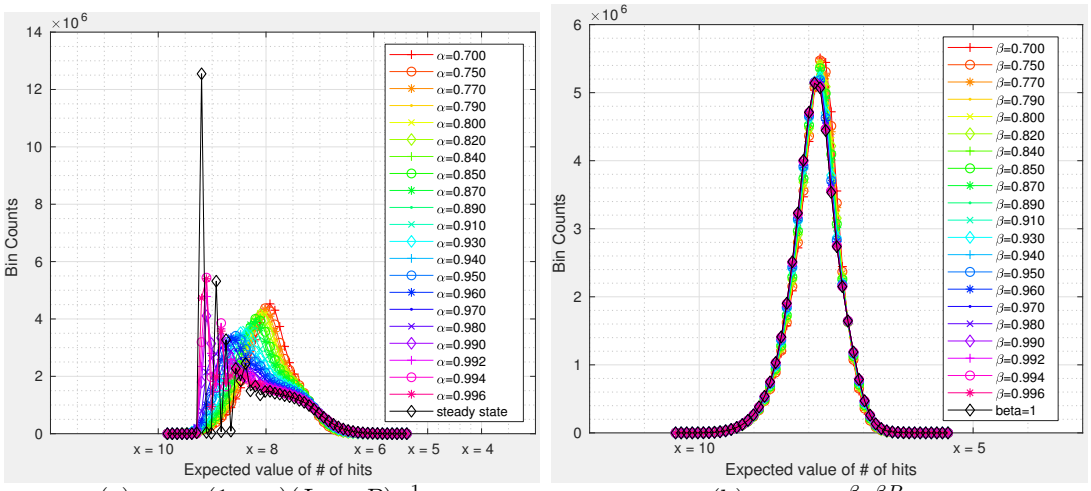
(a) Brin-Page model (b) Chung's model

FIGURE A.14: Comparison in KL-divergence scores between the cases of $\alpha, \beta = 1$ with friendster network



(a) Brin-Page model (b) Chung's model

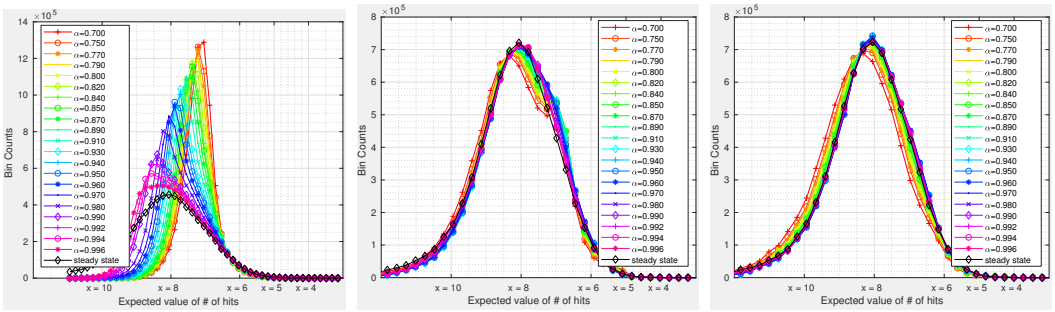
FIGURE A.15: Changing rate of KL-divergence in response to incremental change in damping model parameters with friendster network



(a) $v_\alpha = (1-\alpha)(I-\alpha P)^{-1}v_0$ (b) $v_\beta = e^{-\beta}e^{\beta P}v_0$

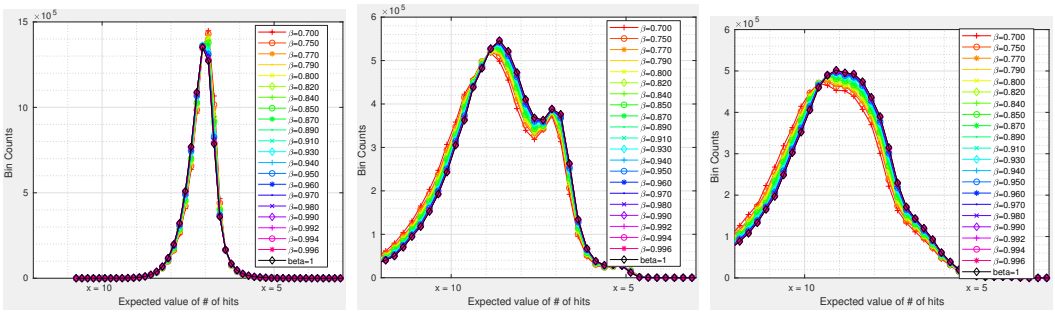
FIGURE A.16: Changes in PageRank distribution with different damping parameter values, with friendster network

A.7.5 Sparsity



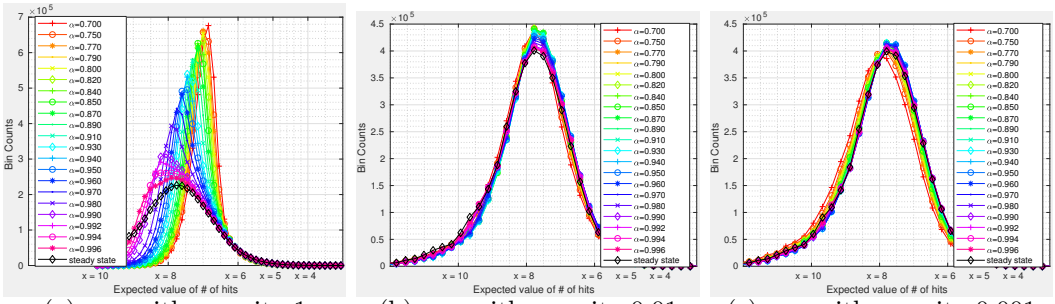
(a) $v_{\alpha 0}$ with sparsity 1 (b) $v_{\alpha 0}$ with sparsity 0.01 (c) $v_{\alpha 0}$ with sparsity 0.001

FIGURE A.17: Changes in PageRank distribution with different sparsity of personalized distribution, with Wikipedia network



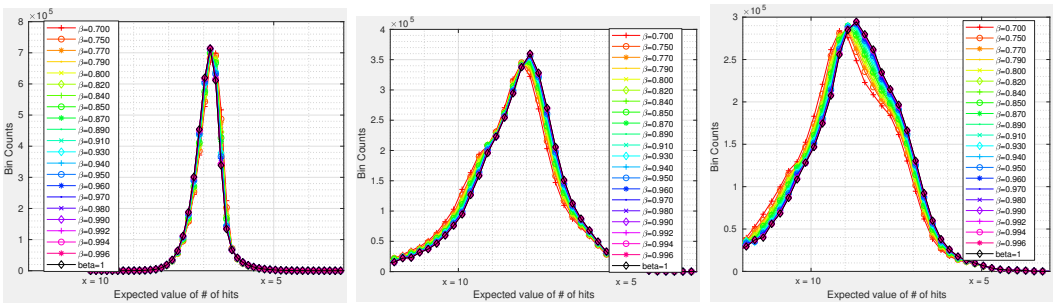
(a) $v_{\beta 0}$ with sparsity 1 (b) $v_{\beta 0}$ with sparsity 0.01 (c) $v_{\beta 0}$ with sparsity 0.001

FIGURE A.18: Changes in PageRank distribution with different sparsity of personalized distribution, with Wikipedia network



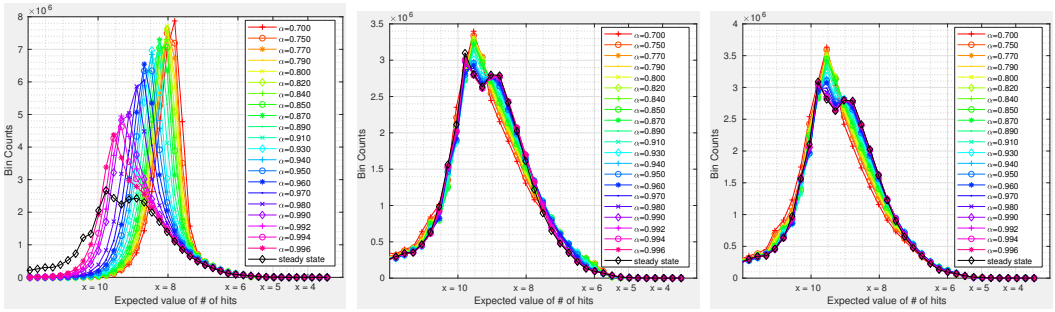
(a) $v_{\alpha 0}$ with sparsity 1 (b) $v_{\alpha 0}$ with sparsity 0.01 (c) $v_{\alpha 0}$ with sparsity 0.001

FIGURE A.19: Changes in PageRank distribution with different sparsity of personalized distribution, with DBpedia network



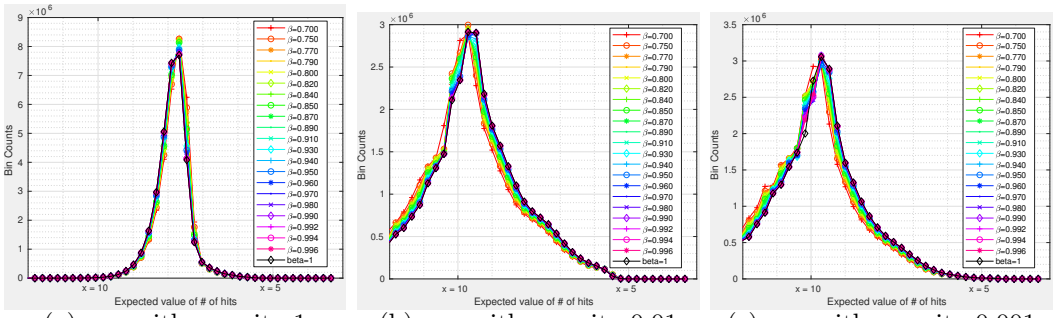
(a) $v_{\beta 0}$ with sparsity 1 (b) $v_{\beta 0}$ with sparsity 0.01 (c) $v_{\beta 0}$ with sparsity 0.001

FIGURE A.20: Changes in PageRank distribution with different sparsity of personalized distribution, with DBpedia network



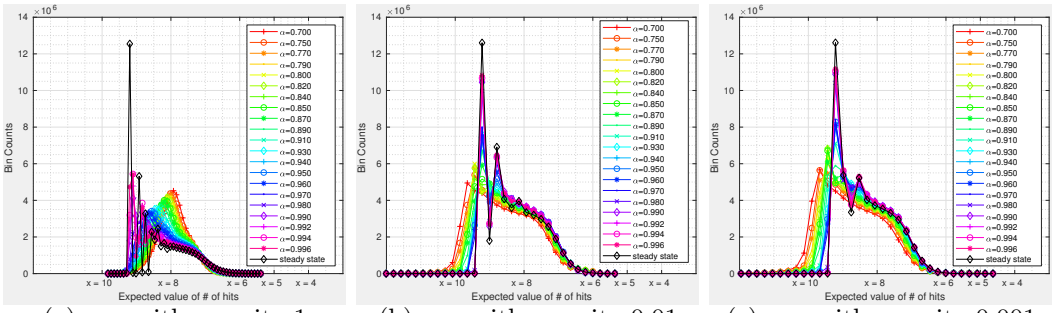
(a) $v_{\alpha 0}$ with sparsity 1 (b) $v_{\alpha 0}$ with sparsity 0.01 (c) $v_{\alpha 0}$ with sparsity 0.001

FIGURE A.21: Changes in PageRank distribution with different sparsity of personalized distribution, with Twitter-www network



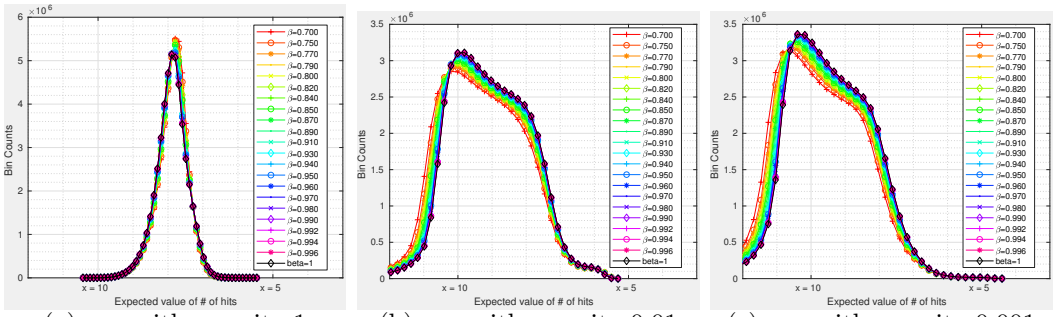
(a) $v_{\beta 0}$ with sparsity 1 (b) $v_{\beta 0}$ with sparsity 0.01 (c) $v_{\beta 0}$ with sparsity 0.001

FIGURE A.22: Changes in PageRank distribution with different sparsity of personalized distribution, with Twitter-www network



(a) $v_{\alpha 0}$ with sparsity 1 (b) $v_{\alpha 0}$ with sparsity 0.01 (c) $v_{\alpha 0}$ with sparsity 0.001

FIGURE A.23: Changes in PageRank distribution with different sparsity of personalized distribution, with friendster network



(a) $v_{\beta 0}$ with sparsity 1 (b) $v_{\beta 0}$ with sparsity 0.01 (c) $v_{\beta 0}$ with sparsity 0.001

FIGURE A.24: Changes in PageRank distribution with different sparsity of personalized distribution, with friendster network

Bibliography

[1] P. Boldi, M. Santini, and S. Vigna. Pagerank as a function of the damping factor. In *Proceedings of the 14th international conference on World Wide Web*, pages 557–566. ACM, 2005.

[2] M. Bressan and E. Peserico. Choose the damping, choose the ranking? *Journal of Discrete Algorithms*, 8(2):199–213, 2010.

[3] S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. In *Seventh International World-Wide Web Conference (WWW 1998)*, 1998.

[4] X. Chen. Exploiting common structures across multiple network propagation schemes. 2018.

[5] X. Chen, Z. Dai, and Y. Qian. Reducible matrix and pagerank with iteration methods. 2017.

[6] F. Chung. The heat kernel as the pagerank of a graph. *Proceedings of the National Academy of Sciences*, 104(50):19735–19740, 2007.

[7] F. Chung. A local graph partitioning algorithm using heat kernel pagerank. *Internet Mathematics*, 6(3):315–330, 2009.

[8] F. Chung and O. Simpson. Solving local linear systems with boundary conditions using heat kernel pagerank. *Internet Mathematics*, 11(4-5):449–471, 2015.

[9] J. W. Demmel. *Applied numerical linear algebra*, volume 56. Siam, 1997.

[10] A. L. Dulmage and N. S. Mendelsohn. Coverings of bipartite graphs. *Canadian Journal of Mathematics*, 10(4):516–534, 1958.

[11] G. H. Golub and C. F. Van Loan. *Matrix computations*, volume 3. JHU Press, 2012.

[12] T. H. Haveliwala. Topic-sensitive pagerank: A context-sensitive ranking algorithm for web search. *IEEE transactions on knowledge and data engineering*, 15(4):784–796, 2003.

[13] G. Jeh and J. Widom. Scaling personalized web search. In *Proceedings of the 12th international conference on World Wide Web*, pages 271–279. Acm, 2003.

[14] S. D. Kamvar, T. H. Haveliwala, C. D. Manning, and G. H. Golub. Extrapolation methods for accelerating pagerank computations. In *Proceedings of the 12th international conference on World Wide Web*, pages 261–270. ACM, 2003.

[15] R. I. Kondor and J. Lafferty. Diffusion kernels on graphs and other discrete input spaces. In *ICML*, volume 2, pages 315–322, 2002.

[16] S. Kullback and R. A. Leibler. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86, 1951.

[17] J. Kunegis. Konect: the koblenz network collection. In *Proceedings of the 22nd International Conference on World Wide Web*, pages 1343–1350. ACM, 2013.

[18] K. S. Miller. On the inverse of the sum of matrices. *Mathematics magazine*, 54(2):67–72, 1981.

[19] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab, 1999.

[20] M. Richardson and P. Domingos. The intelligent surfer: Probabilistic combination of link and content information in pagerank. In *Advances in neural information processing systems*, pages 1441–1448, 2002.