Damping Effect on PageRank Distribution

IEEE High Performance Extreme Computing, Waltham, MA, USA
September 26, 2018

Tiancheng Liu  Yuchen Qian  Xi Chen  Xiaobai Sun

Department of Computer Science, Duke University, USA
Outline

- **Personalized PageRank model:** invention by Brin and Page (1998) in need of innovative extension

- **The PageRank model family:** an analytic apparatus with increased description power and scope

- **Analysis:** damping effects on PageRank distributions

- **Algorithm:** exploiting structures of the personalized, stochastic Krylov (PSK) space

- **Findings:** by experiments on real-world network data
Sparse graphs in sparse matrix representations

link graph \( G(V, E) \)
directed edge \((u, v) \in E\)

adjacency matrix \( A \)
\[ A(v, u) = 1 \]
\( d_{in} \) in-degrees
\( d_{out} \) out-degrees

probability transition matrix \( P \)
\[ P = A \cdot \text{diag}(1./d_{out}) \]
factor form in storage
Web surfing modeled as a random walk on $M_\alpha(v)$, a Markov chain with a \textbf{personalized} term $S$

\begin{align*}
M_\alpha(v) &= \alpha \quad \text{damping factor} \\
&= \alpha \\
&= \alpha \\
= \alpha + (1 - \alpha) S, \\
S &= v \quad \text{personalized vector} \\
&= v \\
&= v \\
&= v \quad e^T \quad \text{gathering vector}
\end{align*}

Bernoulli decision at each click: follow $P$-links or $S$-links with probability $\alpha \in (0, 1)$ a.k.a. \textbf{damping factor}

The personalized term $S$: direct links to $v$-nodes (yellow) gathering/broadcasting rank-1, stochastic
Web surfing modeled as a random walk on $M_{\alpha}(v)$, a Markov chain with a **personalized** term $S$

$$M_{\alpha}(v) = \alpha P + (1 - \alpha) S,$$

where $\alpha$ is the **damping factor**.

![Matrix diagram](image)

$S = v e^T$ where $v$ is the **personalized vector** and $e$ is the **gathering vector**.

Bernoulli decision at each click:
- follow $P$-links or $S$-links
- with probability $\alpha \in (0, 1)$
- a.k.a. **damping factor**

The personalized term $S$:
- direct links to $v$-nodes (yellow)
- gathering/broadcasting
- rank-1, stochastic
Equivalent expressions of PageRank distribution vector

**Purpose:** multi-aspect investigation for *interpretation* and computational analysis

1. Steady state distribution of $M_\alpha$

\[ M_\alpha x = \left[ \alpha P + (1 - \alpha) v e^T \right] x = x \]

the power method

Asymptotic walk on $M_\alpha$, memoryless of $x_0$

2. Solution to sparse linear system

\[ (I - \alpha P) x = (1 - \alpha) v \]

many iterative solution methods

3. Explicit representation

\[ x = (1 - \alpha) \sum_k \alpha^k (P^k v) \]

in Neumann series with $P$, $v$, $\alpha$

Cumulative propagation of $v$ on $P$

4. Differential transition equation

\[ \dot{x}(\alpha) = [P(I - \alpha P)^{-1} - (1 - \alpha)^{-1} I] x(\alpha) \]

spectrum-based method
Outline

- **Personalized PageRank model:**
  invention by Brin and Page (1998)
  in need of innovative extension

- **The PageRank model family:**
  an analytic apparatus with increased
description power and scope

- **Analysis:**
  damping effects on PageRank distributions

- **Algorithm:**
  exploiting structures of the personalized,
  stochastic Krylov (PSK) space

- **Findings:**
  by experiments on real-world network data
PageRank model family: characterizing various propagation patterns

Model description in equivalent expressions:

- **Propagation kernel functions**
  propagation patterns

- **Cumulative propagation on** $P$

- **Linear systems**

- **Differential transitions**
  PageRank distribution response to damping variation

A few particular subfamilies of propagation kernel functions
Propagation kernel functions

**Propagation kernel function** $f_\rho(\lambda)$

$$f_\rho(\lambda) = \sum_k w_k(\rho) \lambda^k$$

PageRank vector (model solution) with particular network $P$ and personalized distribution vector $v$

$$x = f_\rho(P)v = \sum_k w_k(\rho) \cdot P^k v$$

$\{w_k(\rho)\}$: any probability mass function (pmf) of variable $\rho$, w.i./w.o. additional parameters

PageRank distributions of 3 propagation patterns with $P$ for link graph Twitter(www) ¹

¹ H. Kwak et al. (2009)
Conway-Maxwell-Poisson (CMP):

\[ w_k(\rho, \nu) = \frac{\rho^k}{(k!)^\nu Z} \]

Damping speed parameter \( \nu \geq 0 \)

\[ \nu = \begin{cases} 
0, & \text{geometric, (B-P, 1998)} \\
1, & \text{Poisson, (Chung, 2007)} \\
< 1, & \text{slow decaying with } k \\
> 1, & \text{fast decaying with } k 
\end{cases} \]

Slow and fast propagation patterns of CMP distribution

**Slow damping speed:** \( 0 \leq \nu \leq 1 \) \( (\rho = 0.9) \)

including BP model and Chung's model

**Fast damping speed:** \( \nu \geq 1 \) \( (\rho = 5) \)
Propagation pattern kernels: NB sub-family

**Negative Binomial (NB):** step $k$

\[
W_k\left(\rho, \frac{r}{k}\right) = \binom{k + r - 1}{k} \rho^k (1 - \rho)^r
\]

Distribution shape parameter $r$:

\[
r = \begin{cases} 
1, & \text{geometric distribution} \\
\infty, & \text{Poisson distribution, with } r \cdot \frac{\rho}{(1 - \rho)} = \text{const} 
\end{cases}
\]

Propagation patterns of NB distribution
Logarithmic: step $k$

$$w_k(\rho) = \frac{-1}{\ln(1-\rho)} \frac{\rho^k}{k}, \quad \rho \in (0, 1)$$

unique new model in the model family:
weight decay faster than geometric distribution
weight decay slower than Poisson distribution
no extra control parameters
Propagation pattern kernels: precursor models and new model

Precursor models:

Brin-Page\textsuperscript{1} model: \textbf{geometric} distribution

\[ w_k(\alpha) = (1 - \alpha)\alpha^k \]

Chung’s\textsuperscript{2} model: \textbf{Poisson} distribution

\[ w_k(\beta) = e^{-\beta} \frac{\beta^k}{k!} \]

new model in the family:

\textbf{log-\gamma} model: \textbf{logarithmic} distribution

\[ w_k(\gamma) = \frac{-1}{\ln(1 - \gamma)} \frac{\gamma^k}{k} \]

\textsuperscript{1} L. Page and S. Brin, 1998 \textsuperscript{2} F. Chung, PNAS, 2007
Cumulative propagation on $P$ and personalized vector $v$.

- Geometric kernel (Brin-Page): $x(\alpha) = z_\alpha \sum_k \alpha^k p^k v$
- Poisson kernel (Chung): $x(\beta) = z_\beta \sum_k \beta^k p^k v$
- Logarithmic kernel (log-$\gamma$): $x(\gamma) = z_\gamma \sum_k \frac{\gamma^k}{k} p^k v$

Link graph $P$ and propagation on $P$: $v$, $Pv$, $P^2v$, $P^{m-1}v$. 
**Linear systems**

**Close-form** expression of the coefficient matrix

\[ A_\rho(P)x = v, \quad A_\rho(P) = f_\rho^{-1}(P) \]

Particular instances

- **Brin-Page model:**
  \[ A_\alpha(P) = (1 - \alpha)^{-1}(I - \alpha P) \]

- **Chung’s model:**
  \[ A_\beta(P) = e^{-\beta(I - P)} \]

- **log-\(\gamma\) model:**
  \[ A_\gamma(P) = \ln(1 - \gamma) \ln^{-1}(I - \gamma P) \]

- Except the Brin-Page model, explicit formation of the coefficient matrix is non-necessary
- This formulation is used for derivation of the differential transition equation (next)
Differential transition

Effect of damping variation in one model:
Node-wise trajectory of PageRank vector $\dot{x}(\rho)$

$$\dot{x}(\rho) = \frac{d}{d\rho} x(\rho) = \frac{\partial}{\partial \rho} f_\rho(P) v = Q_\rho(P)x(\rho)$$

at any particular value of $\rho$

**Brin-Page model:**

$$Q_\alpha(P) = [P(I - \alpha P)^{-1} - (1 - \alpha)^{-1} I]$$

**Chung’s model:**

$$Q = -(I - P)$$

**log-$\gamma$ model:**

$$Q_\gamma(P) = \frac{(1 - \gamma)^{-1}}{\ln(1 - \gamma)} I - P(I - \gamma P)^{-1} (\ln(I - \gamma P))^{-1}$$

- Matrix-vector multiplication for Chung’s model
- Linear-solver may be used once again for Brin-Page model
- An efficient spectrum-based algorithm for all models, without eigen-decomposition of $P$
Outline

- **Personalized PageRank model:** invention by Brin and Page (1998) in need of innovative extension

- **The PageRank model family:** an analytic apparatus with increased description power and scope

- **Analysis:** damping effects on PageRank distributions

- **Algorithm:** exploiting structures of the personalized, stochastic Krylov (PSK) space

- **Findings:** by experiments on real-world network data
statistically similar damping level of propagation on $P$:
at expected propagation weight center

$$\mu(w_k(\rho)) = \sum_{k \in N_w} k \cdot w_k(\rho)$$

Brin-Page $\leftrightarrow$ Chung’s

$$\frac{\alpha}{1 - \alpha} = \beta$$

Brin-Page $\leftrightarrow$ log-$\gamma$

$$\frac{\alpha}{1 - \alpha} = \left(\frac{\gamma}{1 - \gamma}\right) \frac{-1}{\ln(1 - \gamma)}$$
Intra-model damping effect by KL divergence and its derivative

Aggregated effect of damping variation: KL divergence of PageRank vectors (scalar)

\[
KL(x(\rho), x(\rho_o)) = \sum_{i} x_i(\rho) \log \frac{x_i(\rho)}{x_i(\rho_o)}
\]

\[
\frac{d}{d\rho} KL(x(\rho), x(\rho_o)) = \dot{x}(\rho)(\log x(\rho) - \log x(\rho_o) + e)
\]

* \(dKL/d\rho\) in red, \(KL\) in blue
* reference damping factor denote as \(\rho_0\)
Outline

- **Personalized PageRank model:**
  invention by Brin and Page (1998)
in need of innovative extension

- **The PageRank model family:**
an analytic apparatus with increased description power and scope

- **Analysis:**
damping effects on PageRank distributions

- **Algorithm:**
  exploiting structures of the personalized, stochastic Krylov (PSK) space

- **Findings:**
  by experiments on real-world network data
Personalized, stochastic Krylov space

**Personalized, stochastic Krylov (PSK) space:**
\[
\mathcal{PSK}(P, v) = \text{span}\{v, Pv, P^2v, \ldots, P^kv, \ldots\},
\]
\[v \geq 0, \quad e^Tv = 1\]

**Properties:**
- Any convex combination of the Krylov vectors is a probability distribution
- The same PSK space is shared by all models, housing all model solutions and their trajectories
- The PSK space is of finite dimension \(m\)
- Let \(K = [v, Pv, P^2v, \ldots, P^{m-1}v]\) and \(K = QR\). There exists a Hessenberg matrix \(H\) such that \(PQ = QH\), \(Qe_1 = v\) and that \(g(P)v = Qg(H)e_1\) for any function \(g\)

**PageRank vector**
\[x(\rho) = f_\rho(P)v \in \mathcal{PSK}(P, v)\]

**PageRank vector trajectory**
\[\dot{x}(\rho) = Q_\rho(P)x(\rho) \in \mathcal{PSK}(P, v)\]
Efficient algorithm for damping effect analysis

intra-model, inter-model damping variations, across all models under consideration based on the PSK properties, without eigen-decomposition

\[ \begin{align*}
&\text{Krylov matrix} \\
&\text{QR decomp.} \\
&P_{n \times n} \rightarrow \mathbf{v}_{n \times 1} \\
&K_{n \times m} \\
&Q_{n \times m} \rightarrow Q_{n \times m} \\
&R_{m \times m} \rightarrow H_{m \times m} \\
&\mathbf{g}(P)_{n \times m} = Q_{n \times m} \mathbf{g}(H)_{n \times 1} \\
&PQ_{n \times m} = QH_{n \times m} \\
\end{align*} \]
Outline

- Personalized PageRank model: invention by Brin and Page (1998) in need of innovative extension
- The PageRank model family: an analytic apparatus with increased description power and scope

- Analysis: damping effects on PageRank distributions
- Algorithm: exploiting structures of the personalized, stochastic Krylov (PSK) space
- Findings: by experiments on real-world network data
## Data: real-world large social and knowledge network snapshots

|                  | Total | #nodes | #nodes in LSCC | [max($d_{out}$), $\mu(d_{out})$, max($d_{in}$)] |
|------------------|-------|--------|---------------|-----------------------------------------------|
| Google           | 875,713 | 434,818 | [4209, 8.86, 382] |
| Wikilink         | 12,150,976 | 7,283,915 | [7527, 50.48, 920207] |
| DBpedia          | 18,268,992 | 3,796,073 | [8104, 26.76, 414924] |
| Twitter(www)     | 41,652,230 | 33,479,734 | [2936232, 42.65, 768552] |
| Twitter(mpi)     | 52,579,682 | 40,012,384 | [778191, 47.57, 3438929] |
| Friendster       | 68,349,466 | 48,928,140 | [3124, 32.76, 3124] |

---

1. Google Inc. (2002)  
2. Wikipedia Foundation (2017)  
3. DBpedia (2017)  
4. H. Kwak et al. (2009)  
5. M. Cha et al. (2010)  
6. ArchiveTeam (2011)
Sparse real-world networks under Dulmage-Mendelsohn permutation

Google ($\tau = 8$)  DBpedia ($\tau = 2$)  Wikilink ($\tau = 2$)

Twitter(www) ($\tau = 2$)  Twitter(mpi) ($\tau = 3$)  Friendster ($\tau = 3$)

each point represents a $1000 \times 1000$ block, a block with $\geq \tau$ non-zeros is colored blue
Personalized stochastic Krylov space: small-world phenomenon

Effective $\mathcal{PSK}(P, \nu)$ dimension $m$ by $R_{ii}$ in QR decomposition
Damping effect: KL and $dKL/d\rho$ across models

$\alpha_0 = 0.85$

$\gamma_0 = 0.94146$

$\beta_0 = 5.6$

$\alpha_0 = 0.95$

$\gamma_0 = 0.98831$

$\beta_0 = 19$

substantial different sensitivity patterns across model

B-P model and log-$\gamma$ model are sensitive when damping parameter approaches 1

Chung’s model is less sensitive with damping parameter change, especially with large $\beta$

Twitter(www) dataset
Damping effect: KL and $dKL/d\rho$ across datasets

Google
DBpedia
Wikilink
Twitter(www)
Twitter(mpi)
Friendster

similar trend across 6 datasets

low variation with relatively small $\alpha$

substantially larger variation when $\alpha \to 1$

Brin-Page model, $\alpha_0 = 0.85$
Intra-model variation: PageRank vector profiles across models

Brin-Page model
Chung’s model
log-γ model

PageRank vector profile: normalized histogram of PageRank values
Twitter(www) dataset
Intra-model variation: PageRank vector profiles across datasets

Brin-Page model, $\alpha_0 = 0.85$
Recap

**Intellectual merits**
- **Rich family of PageRank models**
  capturing, differentiating various activities and propagation patterns with quantitative form and speed
- **Unified analysis of damping effects**
  easily instantiated on particular network $P$ and personalized vector $v$
- **The PSK space**
  residence for all model solutions, foundation for efficient model solution methods

**Experimental findings**
- **Model utility**
  inter-model difference in PageRank distribution profile is much greater than intra-model difference
- **Bump/peak in PageRank distribution**
  single, with minority support
- **The PSK dimension**
  with small-world networks, the dimension of personalized, stochastic Krylov space is low, which leads to upper bounds on algorithm complexity
Thank you!

Tiancheng Liu – tcliu [at] cs.duke.edu
Intellectual merits

- Rich family of PageRank models
  capturing, differentiating various activities and propagation patterns with quantitative form and speed
- Unified analysis of damping effects
  easily instantiated on particular network $P$ and personalized vector $v$
- The PSK space
  residence for all model solutions, foundation for efficient model solution methods

Experimental findings

- Model utility
  inter-model difference in PageRank distribution profile is much greater than intra-model difference
- Bump/peak in PageRank distribution
  single, with minority support
- The PSK dimension
  with small-world networks, the dimension of personalized, stochastic Krylov space is low, which leads to upper bounds on algorithm complexity
