Random Surfing on Multipartite Graphs

Athanasios N. Nikolakopoulos, Antonia Korba and John D. Garofalakis

Department of Computer Engineering and Informatics, University of Patras

December 07, 2016

IEEE International Conference on Big Data, IEEE BigData 2016

Outline

Introduction & Motivation

PageRank Model

$$\mathbf{G} = \alpha\mathbf{H} + (1-\alpha)\mathbf{E}$$

Primitivity Adjustment of the Row-Normalized Adjacency Matrix $\mathbf{H}$:

- ***Damping Factor*** α
  - Has received much attention (Generalized Damping Functions (Functional Rankings) [1], Multidamping [5], ...)
- ***Teleportation matrix*** $\mathbf{E}$
  - Little have been done towards its generalization [8].

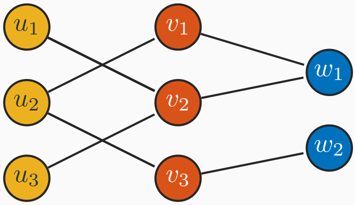
Revisiting the Random Surfer Model II

Problems With Traditional Teleportation

- Treats nodes in a **"leveling way"**
  - Restrictive or even counter-intuitive (eg. structured graphs)
- **Blind to the spectral characteristics** of the underlying graph

Overview of Our Approach

- We focus on **Multipartite Graphs**
- We modify the traditional teleportation model
  - ***Different Teleportation behavior per partite set.***



Block Teleportation Model: Definition, Algorithm and Theoretical Analysis

Model Definition

$$\mathbf{S} = \eta \mathbf{H} + \mu \mathbf{M}$$

$$\mathbf{H} \triangleq \mathrm{diag}(\mathbf{A}_{\mathcal{G}}\mathbf{1})^{-1}\mathbf{A}_{\mathcal{G}},$$

Each partite set is a **Teleportation Block!**

$$[\mathbf{M}]_{ij} \triangleq \begin{cases} \frac{1}{|\mathcal{M}_i|}, & \text{if } v_j \in \mathcal{M}_i, \\ 0, & \text{otherwise,} \end{cases}$$

where $\mathcal{M}_i$ the origin partite set of v_i.

$$\mathbf{M} = \underbrace{\mathbf{R}\boldsymbol{\Delta}\mathbf{R}^{\intercal}}_{\text{Sparse and Low-Rank Factors}}$$

Random Surfing Interpretation

The Random Surfer:

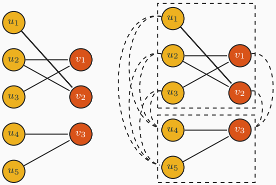
- With probability η follows the edges of the graph
- With probability $\mu \triangleq 1 - \eta$ he jumps to a node belonging to the same partite set he is currently in.

Decomposability and Time-Scale Dissociation

Theorem (Decomposability)

When the value of the teleportation parameter μ is small enough, the Markov chain corresponding to matrix $\mathbf{S}$ is NCD with respect to the partition of the nodes of the initial graph, into different connected components.

$$\mathbf{S} = \tilde{\mathbf{S}} + \varepsilon\mathbf{C}, \qquad \tilde{\mathbf{S}} \triangleq \operatorname{diag}\{\mathbf{S}(\mathcal{G}_1), \ldots, \mathbf{S}(\mathcal{G}_L)\}$$



$$\mathbf{S}^t = \underbrace{\mathbf{Z}_{1_1}}_{\text{Term A}} + \underbrace{\sum_{I=2}^{L} \lambda_{1_I}^t \mathbf{Z}_{1_I}}_{\text{Term B}} + \underbrace{\sum_{I=1}^{L}\sum_{m=2}^{n(I)} \lambda_{m_I}^t \mathbf{Z}_{m_I}}_{\text{Term C}},$$

$$\tilde{\mathbf{S}}^t = \underbrace{\sum_{I=1}^{L} \tilde{\mathbf{Z}}_{1_I}}_{\text{Term }\tilde{\text{A}}} + \underbrace{\sum_{I=1}^{L}\sum_{m=2}^{n(I)} (\tilde{\lambda})_{m_I}^t \tilde{\mathbf{Z}}_{m_I}}_{\text{Term }\tilde{\text{C}}}.$$

- **Short-term Dynamics.**
- **Short-term Equilibrium.**
- **Long-term Dynamics.**
- **Long-term Equilibrium.**

Computational Implications... in brush strokes!

Study each Connected Component in Isolation, and then combine the results.

BT-Rank Algorithm and Computational Analysis

Block-Teleportation Rank

Input: $\mathbf{H}, \mathbf{M} \in \Re^{n \times n}$, ϵ, scalars $\eta, \mu > 0$ such that $\eta + \mu = 1$.

Output: $\boldsymbol{\pi}^\intercal$

1: Let the initial approximation be $\boldsymbol{\pi}^\intercal_{(0)}$. Set $k = 0$.

2: Compute

$$\begin{aligned} \boldsymbol{\pi}^\intercal_{(k+1)} &\leftarrow \boldsymbol{\pi}^\intercal_{(k)} \mathbf{H} \\ \boldsymbol{\phi}^\intercal &\leftarrow \boldsymbol{\pi}^\intercal_{(k)} \mathbf{M} \\ \boldsymbol{\pi}^\intercal_{(k+1)} &\leftarrow \eta \boldsymbol{\pi}^\intercal_{(k+1)} + \mu \boldsymbol{\phi}^\intercal \end{aligned}$$

3: Normalize $\boldsymbol{\pi}^\intercal_{(k+1)}$ and compute $r = \|\boldsymbol{\pi}^\intercal_{(k+1)} - \boldsymbol{\pi}^\intercal_{(k)}\|_1$.

4: If $r < \epsilon$, quit with $\boldsymbol{\pi}^\intercal_{(k+1)}$, otherwise $k \leftarrow k + 1$ and go to step 2.

General Cost: $\boxed{\underbrace{\Theta(\mathrm{nnz}(\mathbf{H}))}_{\text{per iteration}} \frac{\log \epsilon}{\log |\lambda_2(\mathbf{S})|}}$

If $\chi(\mathcal{G}) = 2$ we can dig a little deeper!

Theorem (Eigenvalue Property)

Assuming $\mathcal{G}$ is a connected graph for which $\chi(\mathcal{G}) = 2$ holds, the spectrum of the stochastic matrix $\mathbf{S}$ is such that $-\eta + \mu \in \lambda(\mathbf{S})$.

Theorem (Lumpability)

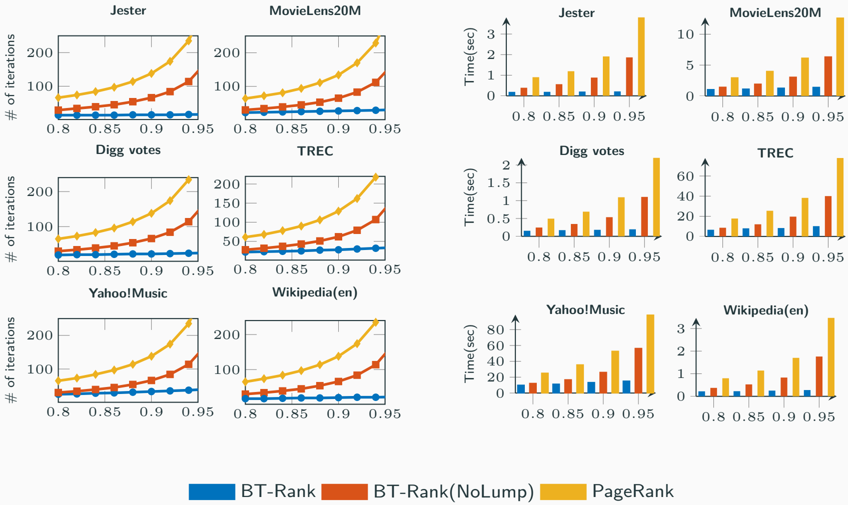
*The BT-Rank Markov chain that corresponds to a 2-chromatic graph, is **lumpable** wrt $\mathcal{A} = \{\mathcal{A}_1, \mathcal{A}_2\}$.*

Theorem (Perron Vector)

When the BT-Rank Markov chain is lumpable wrt to partition $\mathcal{A}$, for the left Perron eigenvector of matrix $\mathbf{S}$ it holds

$$\boldsymbol{\pi}_1^\intercal \mathbf{1}_{\mathcal{A}_1} = \boldsymbol{\pi}_2^\intercal \mathbf{1}_{\mathcal{A}_2}$$

Experimental Evaluation


Jester
MovieLens20M
Digg votes
TREC
Yahoo!Music
Wikipedia(en)
# of iterations
Time(sec)
0.8 0.85 0.9 0.95
BT-Rank
BT-Rank(NoLump)
PageRank

Qualitative Experiments: Ranking-Based Recommendation I

Our Approach

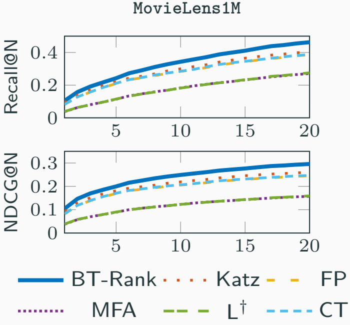
- We model the recommender as a tripartite graph

- Personalization through matrix $\mathbf{M} \triangleq \operatorname{diag}\left(\mathbf{1}\boldsymbol{e}_{\boldsymbol{i}}^{\intercal}, \mathbf{1}\boldsymbol{\omega}_{\boldsymbol{i}}^{\intercal}, \mathbf{1}\boldsymbol{\varpi}_{\boldsymbol{i}}^{\intercal}\right)$
  - $\boldsymbol{\omega}_i$: the normalized vector of the users' ratings over the set of movies.
  - $\boldsymbol{\varpi}_i$: the normalized vector of his mean ratings per genre.

Qualitative Experiments: Ranking-Based Recommendation II

Methodology

- Randomly sample 1.4% of the ratings $\Rightarrow$ probe set $\mathcal{P}$
- Use each item v_j, rated with 5 stars by user u_i in $\mathcal{P}$ $\Rightarrow$ test set $\mathcal{T}$
- Randomly select another 1000 unrated items of the same user for each item in $\mathcal{T}$
- Form ranked lists by ordering all the 1001 items



Metrics

- **Recall@N**
- **Normalized Discounted Cumulative Gain (NDCG@N)**
- **Mean Reciprocal Rank**

Conclusions & Future Work

Conclusion & Future Work

Synopsis

We proposed a simple alternative teleportation component for Random Surfing on Multipartite Graphs.

- Can be handled efficiently
- Entails nice theoretical properties
- Allows for richer modeling

Future Directions

- Propose a Systematic Framework for the definition of teleportation models that match better the underlying graphs
  - For the Web-Graph: **NCDawareRank** (WSDM'13)

References

R. Baeza-Yates, P. Boldi, and C. Castillo.
Generic damping functions for propagating importance in link-based ranking.
Internet Math., 3(4):445–478, 2006.

P.-J. Courtois.
Decomposability: Queueing and Computer System Applications.
ACM monograph series. Academic Press, 1977.

P. Cremonesi, Y. Koren, and R. Turrin.
Performance of recommender algorithms on top-n recommendation tasks.
In *Proceedings of the fourth ACM conference on Recommender systems*, RecSys '10, pages 39–46. ACM, 2010.

F. Fouss, K. Francoisse, L. Yen, A. Pirotte, and M. Saerens.
An experimental investigation of kernels on graphs for collaborative recommendation and semisupervised classification.
Neural Netw., 31:53–72, July 2012.

G. Kollias, E. Gallopoulos, and A. Grama.
Surfing the network for ranking by multidamping.
IEEE Trans. Knowl. Data Eng., 26(9):2323–2336, 2014.

B. Long, X. Wu, Z. M. Zhang, and P. S. Yu.
Unsupervised learning on k-partite graphs.
In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '06, pages 317–326, New York, NY, USA, 2006. ACM.

A. Nikolakopoulos and J. Garofalakis.
NCDREC: A decomposability inspired framework for top-n recommendation.
In *Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2014 IEEE/WIC/ACM International Joint Conferences on*, pages 183–190, Aug 2014.

A. N. Nikolakopoulos and J. D. Garofalakis.
NCDawareRank: a novel ranking method that exploits the decomposable structure of the web.
In *Proceedings of the sixth ACM international conference on Web search and data mining*, WSDM '13, pages 143–152. ACM, 2013.

L. Page, S. Brin, R. Motwani, and T. Winograd.
The pagerank citation ranking: Bringing order to the web.
Technical Report 1999-66, Stanford InfoLab, November 1999.

H. A. Simon and A. Ando.
Aggregation of variables in dynamic systems.
Econometrica, 29(2):pp. 111–138, 1961.

Questions?

Appendix: Eigenvalue Theorem

Back

Theorem (Eigenvalue Property)

Assuming $\mathcal{G}$ is a connected graph for which $\chi(\mathcal{G}) = 2$ holds, the spectrum of the stochastic matrix $\mathbf{S}$ is such that $-\eta + \mu \in \lambda(\mathbf{S})$.

Proof Sketch.

- We define a vector $\mathbf{v} \triangleq [\ \overbrace{1\ \ 1\ \ 1\ \ \cdots\ \ 1}^{\#\text{nodes of 1st partite set}}\quad \overbrace{-1\ \ -1\ \ \cdots\ \ -1}^{\#\text{nodes of 2nd partite set}}\ ]$
- We show that $(-1, \mathbf{v})$ is an eigenpair of matrix $\mathbf{H}$, and $(1, \mathbf{v})$ is an eigenpair of matrix $\mathbf{M}$.
- Then, using any nonsingular matrix, $\mathbf{Q} \triangleq \begin{pmatrix} \mathbf{1} & \mathbf{v} & \mathbf{X} \end{pmatrix}$, allows us to perform a similarity transformation

$$\begin{aligned} \mathbf{Q}^{-1}\mathbf{S}\mathbf{Q} &= \mathbf{Q}^{-1}\left(\eta\mathbf{H} + \mu\mathbf{M}\right)\mathbf{Q} = \cdots = \\ &= \begin{pmatrix} 1 & 0 & \eta\mathbf{y_1}^\intercal\mathbf{H}\mathbf{X} + \mu\mathbf{y_1}^\intercal\mathbf{M}\mathbf{X} \\ 0 & -\eta + \mu & \eta\mathbf{y_2}^\intercal\mathbf{H}\mathbf{X} + \mu\mathbf{y_2}^\intercal\mathbf{M}\mathbf{X} \\ \mathbf{0} & \mathbf{0} & \eta\mathbf{Y}^\intercal\mathbf{H}\mathbf{X} + \mu\mathbf{Y}^\intercal\mathbf{M}\mathbf{X} \end{pmatrix} \end{aligned} \tag{1}$$

that reveals the desired eigenvalue.

□

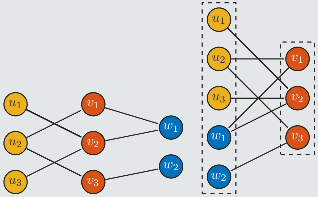
Appendix: Lumpability Theorem

Back

Theorem (Lumpability of the BT-Rank Chain)

The BT-Rank Markov chain that corresponds to a 2-chromatic graph, is ***lumpable****.*

Proof Sketch.



$$\mathbf{S} = \left(\begin{array}{ccccc:ccc}
\frac{\mu}{3} & \frac{\mu}{3} & \frac{\mu}{3} & & & & \eta & \\
\frac{\mu}{3} & \frac{\mu}{3} & \frac{\mu}{3} & & & \frac{\eta}{2} & & \frac{\eta}{2} \\
\frac{\mu}{3} & \frac{\mu}{3} & \frac{\mu}{3} & & & & \eta & \\
 & & & \frac{\mu}{2} & \frac{\mu}{2} & \frac{\eta}{2} & \frac{\eta}{2} & \\
 & & & \frac{\mu}{2} & \frac{\mu}{2} & & & \eta \\
\hdashline
 & \frac{\eta}{2} & & \frac{\eta}{2} & & \frac{\mu}{3} & \frac{\mu}{3} & \frac{\mu}{3} \\
 & & \frac{\eta}{2} & \frac{\eta}{2} & & \frac{\mu}{3} & \frac{\mu}{3} & \frac{\mu}{3} \\
 & \frac{\eta}{2} & & & \frac{\eta}{2} & \frac{\mu}{3} & \frac{\mu}{3} & \frac{\mu}{3}
\end{array}\right)$$

We have

- $\Pr\{i \to \mathcal{A}_2\} = \sum_{j \in \mathcal{A}_2} S_{ij} = \eta$ for all $i \in \mathcal{A}_1$.
- $\Pr\{i \to \mathcal{A}_1\} = \sum_{j \in \mathcal{A}_1} S_{ij} = \eta$ for all $i \in \mathcal{A}_2$.

$\square$

Appendix: Perron Vector Theorem

Back

Theorem (Eigenvector Property)

When the BT-Rank Markov chain is lumpable with respect to partition $\mathcal{A} = \{\mathcal{A}_1, \mathcal{A}_2\}$, for the left Perron-Frobenius eigenvector of matrix $\mathbf{S}$ it holds

$$\boldsymbol{\pi}_1^\intercal \mathbf{1}_{\mathcal{A}_1} = \boldsymbol{\pi}_2^\intercal \mathbf{1}_{\mathcal{A}_2}$$

Proof Sketch.

$$\begin{aligned}
\boldsymbol{\pi}^\intercal \mathbf{S} \begin{pmatrix} \mathbf{1}_{\mathcal{A}_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{\mathcal{A}_2} \end{pmatrix} &= \boldsymbol{\pi}^\intercal \begin{pmatrix} \mathbf{1}_{\mathcal{A}_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{\mathcal{A}_2} \end{pmatrix} \\
\boldsymbol{\pi}^\intercal \begin{pmatrix} \mu \mathbf{M}_{11} \mathbf{1}_{\mathcal{A}_1} & \eta \mathbf{H}_{12} \mathbf{1}_{\mathcal{A}_2} \\ \eta \mathbf{H}_{21} \mathbf{1}_{\mathcal{A}_1} & \mu \mathbf{M}_{22} \mathbf{1}_{\mathcal{A}_2} \end{pmatrix} &= \begin{pmatrix} \boldsymbol{\pi}_1^\intercal \mathbf{1}_{\mathcal{A}_1} & \boldsymbol{\pi}_2^\intercal \mathbf{1}_{\mathcal{A}_2} \end{pmatrix} \\
\begin{pmatrix} \boldsymbol{\pi}_1^\intercal & \boldsymbol{\pi}_2^\intercal \end{pmatrix} \begin{pmatrix} \mu \mathbf{1}_{\mathcal{A}_1} & \eta \mathbf{1}_{\mathcal{A}_1} \\ \eta \mathbf{1}_{\mathcal{A}_2} & \mu \mathbf{1}_{\mathcal{A}_2} \end{pmatrix} &= \begin{pmatrix} \boldsymbol{\pi}_1^\intercal \mathbf{1}_{\mathcal{A}_1} & \boldsymbol{\pi}_2^\intercal \mathbf{1}_{\mathcal{A}_2} \end{pmatrix}
\end{aligned}$$

and the result follows from the solution of the system. $\square$

Appendix: Near Complete Decomposability

Back

Herbert A. Simon

Sparsity $\longleftrightarrow$ Hierarchy $\longleftrightarrow$ Decomposability [10].

Nearly Completely Decomposable Systems

Interactions: Strong Within Blocks - Weak Between Blocks

- Successful Applications in Diverse Disciplines
  - Economics, Cognitive Theory, Management, Biology, Ecology, etc.
- Computer Science and Engineering:
  - Performance Analysis (Courtois [2])
  - Web Search (NG13 [8])
  - Recommendation and Data Mining (NG14 [7])