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
1. Introduction & Motivation

2. Block Teleportation Model: Definition, Algorithm and Theoretical Analysis

3. Experimental Evaluation

4. Conclusions & Future Work
Introduction & Motivation
PageRank Model

\[ G = \alpha H + (1 - \alpha)E \]

Primitivity Adjustment of the Row-Normalized Adjacency Matrix \( H \):

- **Damping Factor** \( \alpha \)
  - Has received much attention (Generalized Damping Functions (Functional Rankings) [1], Multidamping [5], ...)

- **Teleportation matrix** \( E \)
  - 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 (e.g. 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

\[ S = \eta H + \mu M \]

\[ H \triangleq \text{diag}(A_G 1)^{-1} A_G, \]

Each partite set is a Teleportation Block!

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

where \( M_i \) the origin partite set of \( v_i \).

Random Surfing Interpretation

The Random Surfer:

- With probability \( \eta \) 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 $\mu$ is small enough, the Markov chain corresponding to matrix $S$ is NCD with respect to the partition of the nodes of the initial graph, into different connected components.

$$S = \tilde{S} + \varepsilon C, \quad \tilde{S} \triangleq \text{diag}\{S(G_1), \ldots, S(G_L)\}$$

\[ S^t = \underbrace{Z_{11}}_{\text{Term A}} + \sum_{I=2}^{L} \lambda^t_{1I} Z_{1I} + \sum_{I=1}^{L} \sum_{m=2}^{n(I)} \lambda^t_{mI} Z_{mI}, \]

\[ \tilde{S}^t = \sum_{I=1}^{L} \tilde{Z}_{1I} + \sum_{I=1}^{L} \sum_{m=2}^{n(I)} (\tilde{\lambda})^t_{mI} \tilde{Z}_{mI}. \]

- **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.
**Input:** \( H, M \in \mathbb{R}^{n \times n}, \epsilon, \text{ scalars} \)
\( \eta, \mu > 0 \text{ such that } \eta + \mu = 1. \)

**Output:** \( \pi^T \)
1. Let the initial approximation be \( \pi^T_{(0)} \). Set \( k = 0. \)
2. Compute
   \[
   \begin{align*}
   \pi^T_{(k+1)} & \leftarrow \pi^T_{(k)} H \\
   \phi^T & \leftarrow \pi^T_{(k)} M \\
   \pi^T_{(k+1)} & \leftarrow \eta \pi^T_{(k+1)} + \mu \phi^T
   \end{align*}
   \]
3. Normalize \( \pi^T_{(k+1)} \) and compute
   \[
   r = \| \pi^T_{(k+1)} - \pi^T_{(k)} \|_1.
   \]
4. If \( r < \epsilon \), quit with \( \pi^T_{(k+1)} \), otherwise \( k \leftarrow k + 1 \) and go to step 2.

**General Cost:**
\[
\Theta(\text{nnz}(H)) \frac{\log \epsilon}{\log |\lambda_2(S)|}
\]

**If** \( \chi(G) = 2 \) **we can dig a little deeper!**

**Theorem (Eigenvalue Property )**

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

**Theorem (Lumpability )**

The BT-Rank Markov chain that corresponds to a 2-chromatic graph, is lumpable wrt \( A = \{A_1, A_2\} \).

**Theorem (Perron Vector )**

When the BT-Rank Markov chain is lumpable wrt to partition \( A \), for the left Perron eigenvector of matrix \( S \) it holds
\[
\pi^T_1 \mathbf{1}_{A_1} = \pi^T_2 \mathbf{1}_{A_2}
\]
Experimental Evaluation
Computational Experiments

![Graphs showing computational experiments with various datasets including Jester, MovieLens20M, Digg votes, TREC, Yahoo!Music, and Wikipedia(en). The graphs display the number of iterations and time taken for different algorithms: BT-Rank, BT-Rank(NoLump), and PageRank.]
Our Approach

- We model the recommender as a tripartite graph

- Personalization through matrix $\mathbf{M} \triangleq \text{diag} \left( \mathbf{1} e_i^T, \mathbf{1} \omega_i^T, \mathbf{1} \varpi_i^T \right)$
  - $\omega_i$: the normalized vector of the users’ ratings over the set of movies.
  - $\varpi_i$: the normalized vector of his mean ratings per genre.
Methodology

- Randomly sample 1.4% of the ratings ⇒ probe set \( \mathcal{P} \)
- Use each item \( v_j \), rated with 5 stars by user \( u_i \) in \( \mathcal{P} \) ⇒ 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

![Graph showing Recall@N, NDCG@N, and MRR for different methods: BT-Rank, Katz, FP, MFA, L†, CT]
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)
Generic damping functions for propagating importance in link-based ranking.

P.-J. Courtois.
*Decomposability: Queueing and Computer System Applications.*

P. Cremonesi, Y. Koren, and R. Turrin.
Performance of recommender algorithms on top-n recommendation tasks.

An experimental investigation of kernels on graphs for collaborative recommendation and semisupervised classification.

G. Kollias, E. Gallopoulos, and A. Grama.
Surfing the network for ranking by multidamping.

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.

The pagerank citation ranking: Bringing order to the web.

H. A. Simon and A. Ando.
Aggregation of variables in dynamic systems.
Questions?
Appendix: Eigenvalue Theorem

**Theorem (Eigenvalue Property)**

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

**Proof Sketch.**

- We define a vector $v \triangleq \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & -1 & -1 & \cdots & -1 \end{bmatrix}$
- We show that $(-1, v)$ is an eigenpair of matrix $H$, and $(1, v)$ is an eigenpair of matrix $M$.
- Then, using any nonsingular matrix, $Q \triangleq \begin{bmatrix} 1 & v & X \end{bmatrix}$, allows us to perform a similarity transformation

$$
Q^{-1}SQ = Q^{-1} (\eta H + \mu M) Q = \cdots =
\begin{pmatrix}
1 & 0 & \eta y_1^t HX + \mu y_1^t MX \\
0 & -\eta + \mu & \eta y_2^t HX + \mu y_2^t MX \\
0 & 0 & \eta Y^t HX + \mu Y^t MX
\end{pmatrix}
$$

(1)

that reveals the desired eigenvalue.
Theorem (Lumpability of the BT-Rank Chain)

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

**Proof Sketch.**

We have

- \( \Pr\{i \to A_2\} = \sum_{j \in A_2} S_{ij} = \eta \) for all \( i \in A_1 \).
- \( \Pr\{i \to A_1\} = \sum_{j \in A_1} S_{ij} = \eta \) for all \( i \in A_2 \).
Theorem (Eigenvector Property)
When the BT-Rank Markov chain is lumpable with respect to partition $A = \{A_1, A_2\}$, for the left Perron-Frobenius eigenvector of matrix $S$ it holds

$$\pi_1^T 1_{A_1} = \pi_2^T 1_{A_2}$$

Proof Sketch.

$$\pi^T S \begin{pmatrix} 1_{A_1} & 0 \\ 0 & 1_{A_2} \end{pmatrix} = \pi^T \begin{pmatrix} 1_{A_1} & 0 \\ 0 & 1_{A_2} \end{pmatrix}$$

$$\pi^T \begin{pmatrix} \mu M_{11} 1_{A_1} & \eta H_{12} 1_{A_2} \\ \eta H_{21} 1_{A_1} & \mu M_{22} 1_{A_2} \end{pmatrix} = \begin{pmatrix} \pi_1^T 1_{A_1} & \pi_2^T 1_{A_2} \end{pmatrix}$$

$$\begin{pmatrix} \pi_1^T \\ \pi_2^T \end{pmatrix} \begin{pmatrix} \mu 1_{A_1} & \eta 1_{A_1} \\ \eta 1_{A_2} & \mu 1_{A_2} \end{pmatrix} = \begin{pmatrix} \pi_1^T 1_{A_1} & \pi_2^T 1_{A_2} \end{pmatrix}$$

and the result follows from the solution of the system.
Herbert A. Simon

Sparsity $\leftrightarrow$ Hierarchy $\leftrightarrow$ 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])