Using Local Structures for Network Analysis and Optimization

Prof. Dr. Karsten Weihe
Dr. Lachezar Krumov

Fachbereich Informatik
Fachgebiet Algorithmik
Technische Universität Darmstadt
Overview

- Research Directions
- Motivation
- Introduction to Network Motifs
- Co-Authorship Networks
- Discussion
Research Directions

- Improved time-table information for the Deutsche Bahn
  - Pareto optimal time-table information
  - Updated time-table information under delays
- Analysis of complex networks
  - Co-Authorship networks
- Optimization of distributed systems
  - Structured Peer-to-Peer (P2P) networks
  - P2P live-streaming networks
  - Multi-hop networks
  - P2P Overlays
Motivation

Complex Network
Motivation
Motivation

Topology

Complex Network

Dynamic Function
Motivation

- Complex Network
- Topology
- Dynamic Function
- Local Structures
Introduction to Network Motifs


- Induced subgraphs of 3-4 nodes
- Reflect the local environment of the nodes
- Local perspective beyond the scope of a single node
- Used exclusively as a statistical tool to measure deviations
Co-Authorship Networks

- Analyze the local structures in co-authorship networks
- Nodes represent authors, two authors are connected if they have ever published together
Co-Authorship Networks

- Analyze the local structures in co-authorship networks
- Nodes represent authors, two authors are connected if they have ever published together
Co-Authorship Networks: Publication Data

▪ Question: Are some collaboration patterns more successful than others?
▪ Success: Average number of citations by other publications

▪ Two large publication databases: DBLP and CiteSeerX
▪ Citation frequencies acquired from CiteSeerX and Google Scollar by using 107 self-implemented WebCrawlers
Citation Frequencies as Edge Weights

- Analyze the distribution of citation frequencies among the diff. motifs
- Project the citation frequencies as edge weights
- Four different ways to do that:
  - Directly
  - Divide by the number of mutual publications
  - Divide by the number of involved authors
  - Divide by both, the number of authors and the number of mutual publications
High Citation Frequencies: the Box Motif

- DBLP on the lefthand side, CiteSeerX on the righthand side
- Compared to a shuffled-weights Null-Model
High Citation Frequencies: the Box Motif

- Average motif weights over the past 20 years (DBLP)
- In all 20 cases the box motif prevails again
- Further observations: change in weight of other motifs over the years
The Box Motif: A Deeper Look

- Example of a typical box motif
- The topology of the box motif implies a certain degree of separation
The Box Motif: Segregation in Rank

- Sort the box motif instances according to their weight or top author
- Count the instances where the two heavy authors are directly connected
- In the heavy box motif instances, the two heavy authors are directly connected
- Separation in rank
The Box Motif: Segregation in Time

- Compute the construction time for each motif instance
- Time between the first and the last publication that constitute the motif instance
- Compute the average motif weight for each construction time bin
- Long construction times lead to higher success

Diagram: Graph showing the average edge weight over construction time in years for different motifs.
The Box Motif: Segregation in Discipline

- Compute the average edge betweenness of different motifs
- The box motif constitutes high betweenness values
Co-Authorship Networks: Conclusions

- Projected citation frequencies on Co-authorship networks
- The box motif has unexpectedly high citation frequency per motif edge
- Over two large databases, 4 normalization strategies & the past 20 years
- The box motif edges show a certain degree of segregation in:
  - Rank
  - Time
  - Discipline
- Motifs reveal non-trivial correlation between topology and function
Discussion

- New methodology for distributed control
  - Steer network properties through their local structures
  - Fast, straightforward and easy to apply

- Prominent examples of complex networks:
  - The Web
  - Many web-based networks
    - Online social networks
    - Peer forums
    - Wikipedia
Discussion

- Explore the Web and web-based networks from two perspectives

- Static perspective:
  - Analyze their local structures
  - Reveal the processes shaping those networks

- Active perspective
  - Use local structures to improve web-based systems
  - “Smarter” automatic content detection
  - More accurate recommendation systems
  - More precise text-processing approaches
Thank you very much for your attention!

- [1] Lachezar Krumov, Christoph Fretter, Matthias Müller-Hannemann, Karsten Weihe, Marc-Thorsten Hütt

- [2] Dirk Bradler, Lachezar Krumov, Max Mühlhäuser and Jussi Kangasharju


- [4] Lachezar Krumov, Immanuel Schweizer, Dirk Bradler and Thorsten Strufe
  "Leveraging Network Motifs for the Adaption of Structured Peer-to-Peer Networks", GlobeCom 2010.

- [5] Lachezar Krumov, Adriana Andreeva and Thorsten Stufe

- [6] Christoph Fretter, Lachezar Krumov, Karsten Weihe, Matthias Müller-Hannemann, Marc-Thorsten Hütt

- [7] Dirk Bradler, Lachezar Krumov, Michael Wagner and Jussi Kangasharju

- [8] Lachezar Krumov