Aalto University
School of Science

Networks over time

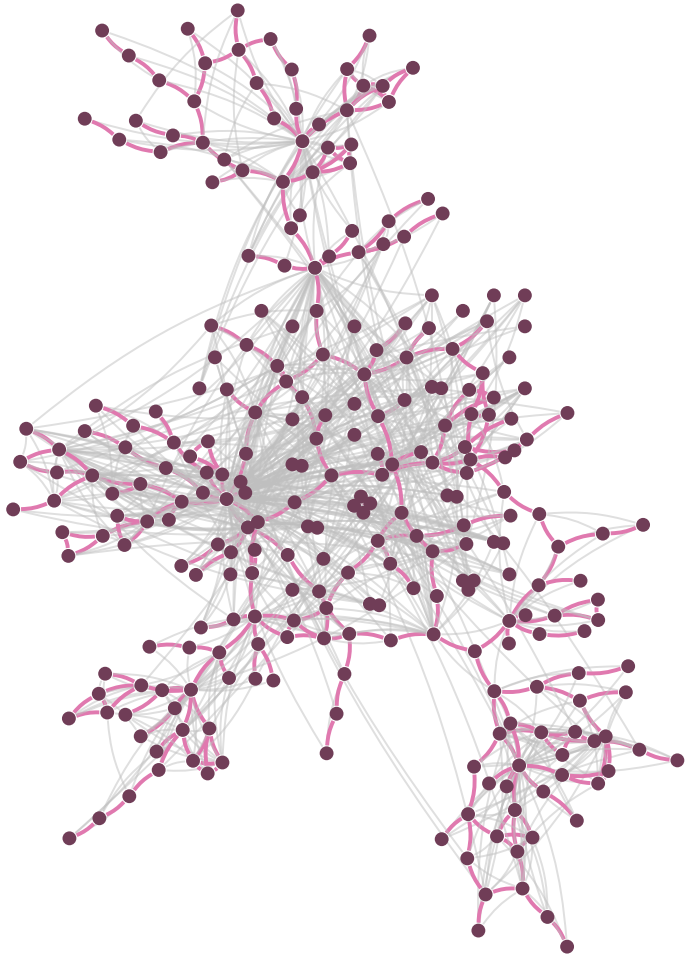
Aristides Gionis

Computer science research day

Oct 12, 2017

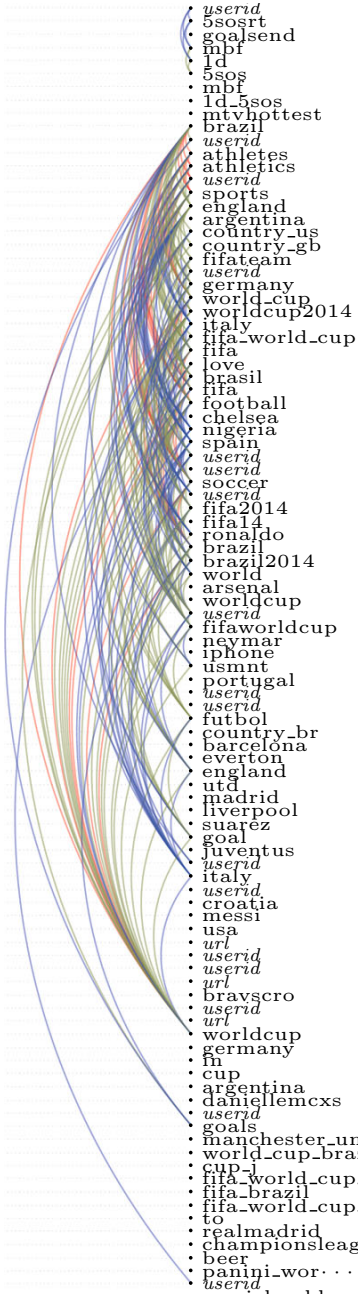
interconnected world

- networks model objects and their relations
- many different network types
  - social
  - informational
  - technological
  - biological
  - …



transforming our society

- online communication networks and social media
- profound implications in
  - knowledge creation
  - information sharing
  - education
  - democracy
  - society as a whole



research questions

- structure discovery
  - finding communities, events, roles of individuals
- study complex dynamic phenomena
  - evolution, information diffusion, opinion formation
- develop novel applications
- design efficient algorithms

some research projects presented today

- polarization in social networks

  [Kiran Garimella, Antonis Matakos]

  - can we identify / measure polarized discussions?
  - can we reduce / moderate polarization?

- AncestryAI [Eric Malmi]

  - reconstruct family trees by linking historical records
  - analyze the resulting family trees

- semantic homophily in online communication

  [Sanja Šćepanović]

  - how to measure homophily in social networks?

networks over time

traditional view

- networks represented as pure graph-theory objects

  no additional vertex / edge information

- emphasis on static networks

- dynamic settings model structural changes

  vertex / edge additions / deletions

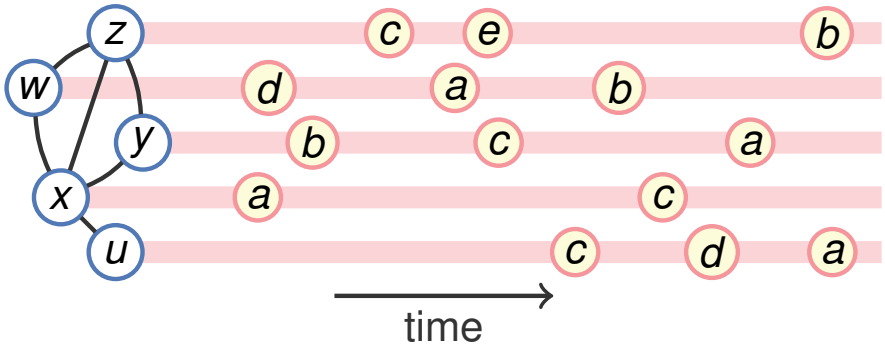
networks over time

modern view

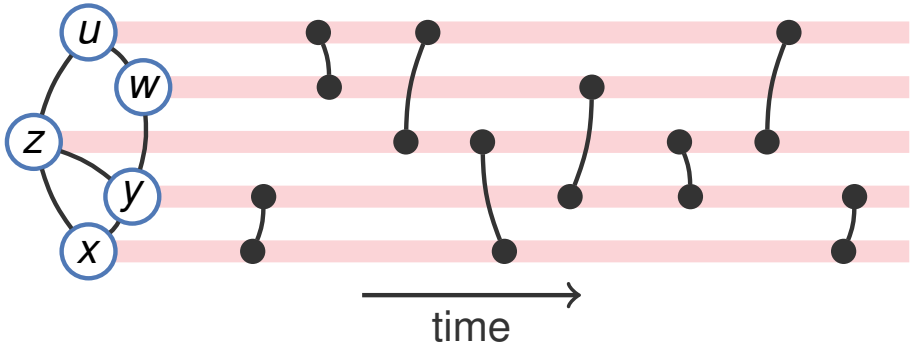
- ability to collect and store large volumes of network data
- available data have fine granularity
- network topology is relatively stable, while lots of activity and interaction is taking place
- giving rise to new concepts, new problems, and new computational challenges

modeling activity in networks

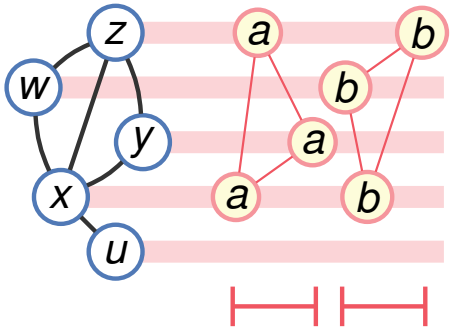
1. network nodes perform actions (e.g., posting messages)



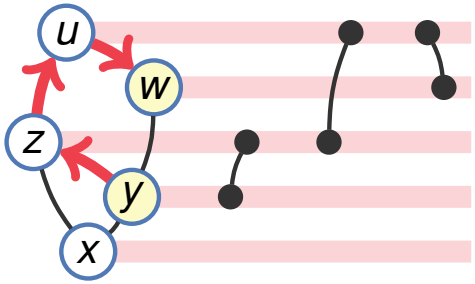
2. network nodes interact with each other
   (e.g., a “like”, a repost, or sending a message to each other)



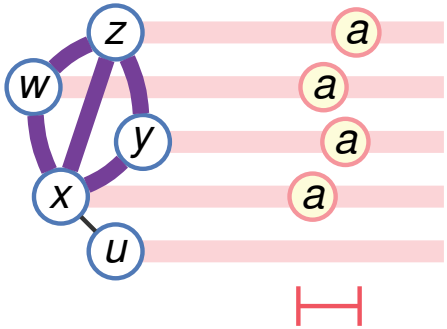
many novel and interesting concepts



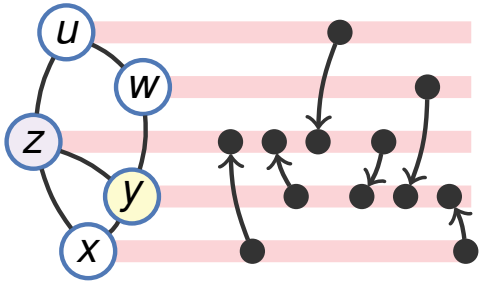
new pattern types



temporal information paths



new types of events



network evolution

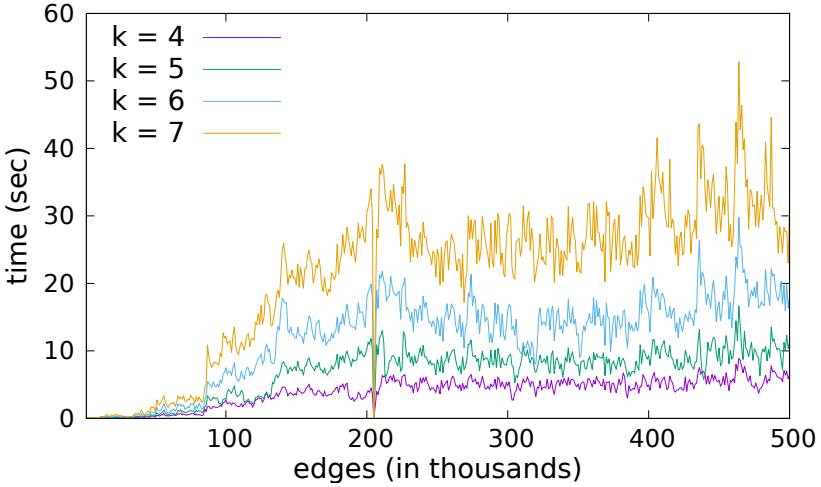
mining temporal networks

- many new interesting problems [Polina Rozenshtein]
- e.g., tracking important nodes
  - maintaining neighborhood profiles
  - temporal PageRank

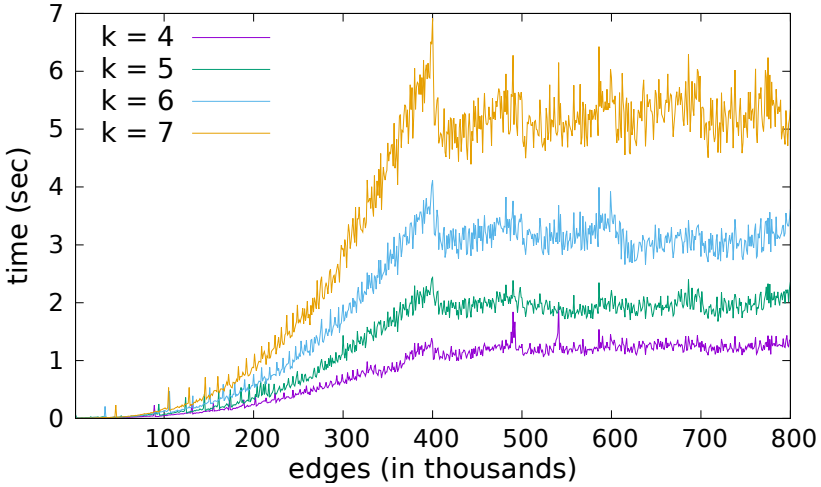
maintaining neighborhood profiles

- problem: for all nodes in the network, track the number of neighbors at distance d
- exact solution requires all-pairs shortest path computation
  - non scalable
- resort to approximations based on diffusion methods
  - approximate memory-efficient streaming algorithm

empirical evaluation — running time



(c) Higgs



(d) DBLP

performance on DBLP dataset

- offline HyperANF : 3.6 sec / sliding window
- proposed approach : 0.003 sec / sliding window

PageRank

- classic approach for measuring node importance
- listed in the top-10 most important data-mining algorithms [Wu et al., 2008]
- numerous applications
  - ranking web pages
  - trust and distrust computation
  - finding experts in social networks
  - ...

research questions and objectives

- extend PageRank to incorporate temporal information and network dynamics
- adapt PageRank to reflect changes in network dynamics and node importance
- estimate importance of a node u at any given time t

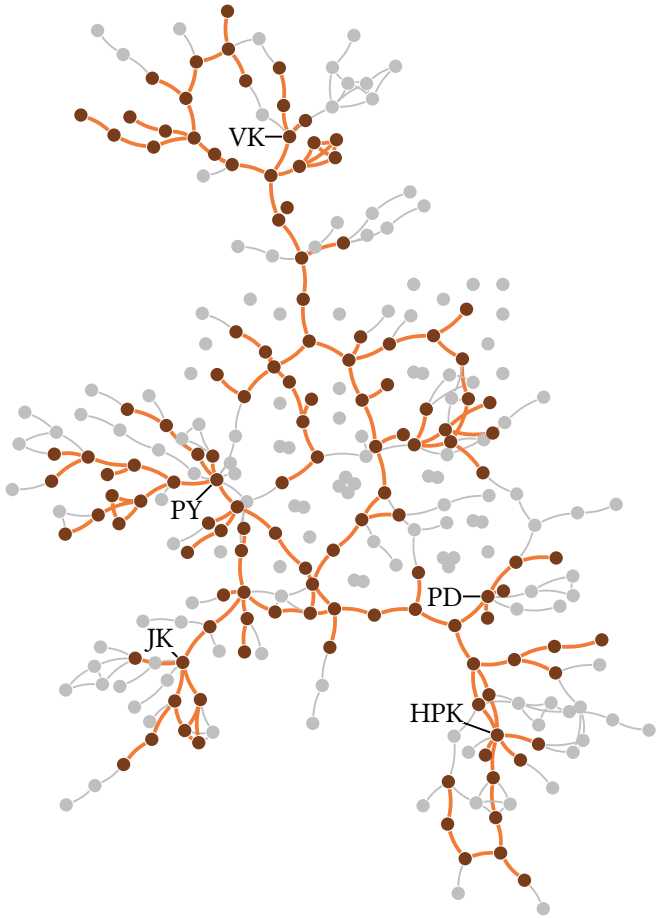
research questions and objectives

- extend PageRank to incorporate temporal information and network dynamics
- adapt PageRank to reflect changes in network dynamics and node importance
- estimate importance of a node u at any given time t
- our solution: new temporal PageRank model
  efficient streaming algorithm

summary

- examples of mining temporal networks
  - maintaining sliding-window neighborhood profiles
  - temporal PageRank
- potential for new concepts, new problem definitions, new computational methods, and new applications

thank you



references

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Philip, S. Y., et al. (2008).

Top 10 algorithms in data mining.

KAIS.