Uncovering Multi-Modal Spread Modes using Joint Diagonalization in Dynamic Human Contact Networks

Damien Fay, Jérôme Kunegis, Eiko Yoneki

University Cambridge, University of Koblenz-Landau, University of Cambridge

Email: eiko.yoneki@cl.cam.ac.uk
http://www.cl.cam.ac.uk/~ey204

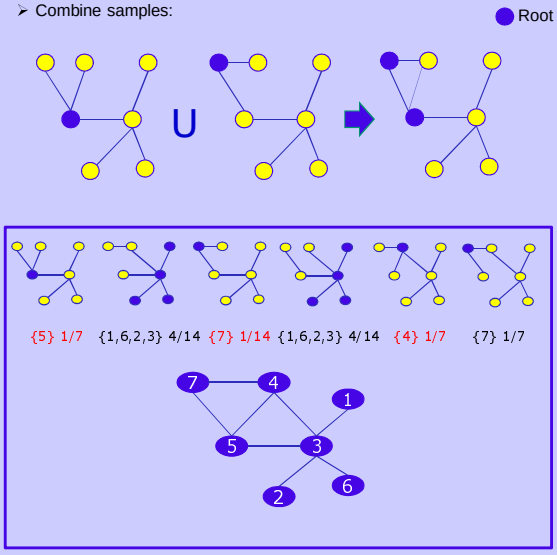
UNIVERSITY OF CAMBRIDGE

CENTRALITY MEASURES

- Think about spread of messages, infectious disease, and gossip
- Understand flow and control: important to uncover centrality nodes
- Most centrality measures are dependent on the number of spanning trees that include a link
  - High link centrality ≡ many spanning trees pass through this link
- Implicit assumption
  - Every spanning tree is equally likely
  - Real network generates traffic proportional to spanning trees

SHORTEST PATH SAMPLING OF SPANNING TREES

- Combine samples:



- Notation: {possible root nodes} probability.
- Example: Spanning tree 1: {5} 1/7 – may begin at 5 only, probability of observing this tree = 1/7
- Distribution of sampling trees is severely skewed

CONTACT NETWORKS

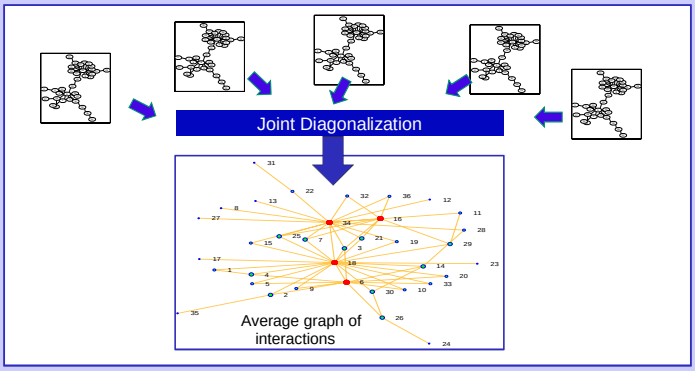
- A Contact network: real-world network of time dependent contacts
  - A → B → C → A
- Problem: How do we uncover modes of spread in multiple groups?
- Typical approach: Cluster nodes to build single network or multiple networks within the sliced time windows
  - Aggregate the number of contacts
  - Ignores time
  - Ignores correlation between links
- Solution: Use spanning tree based samples of a network
  - Akin to spreading a disease in the population and recording the order of infection
  - Order of infection retains time and contact information (i.e. link appearance correlation).
  - Define an eigen-space average across these trees
  - The distribution of deviations gives the required groups

APPLY JOINT DIAGONALIZATION

- Combine many of sampled trees using joint diagonalization
- Ex. Cambridge student contact data: 36 users; 2 weeks; 60 second intervals
- Eigenvalue decomposition of a matrix may be expressed as: $A = U\Lambda U^T$
  - Where, eigenvectors are a special basis relative to A such that Λ is diagonal
- Given N square matrices $A_1, \ldots, A_N$ we seek a basis such that: $A_i = UC_iU^T$
  - where the off-diagonal elements of the projections are minimised
- **Find a basis giving an average eigen-structure of the matrices**

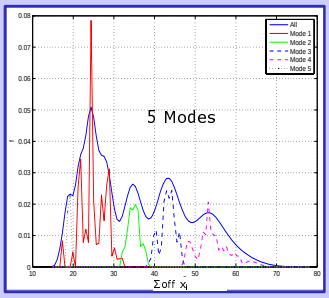
DEVIATION and JOINT DIAGONALIZATION

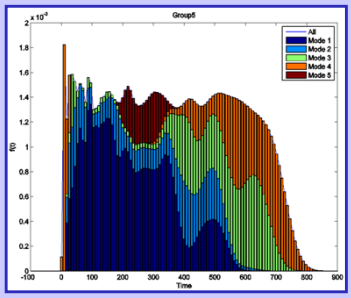
- Simply define the deviation from the average eigen-space as the sum of off-diagonal elements
- Each eigenvalue divides the data set in two equally weighted groups
  => hierarchical segmentation
- Calculate the deviation from the average → **Average Graph of Interaction**



EXAMINE EMPIRICAL DISTRIBUTION of DEVIATION VALUE

- **Multi-modal: each mode represents a different behaviour in the contact network**
- Examine the distribution of modes over time
- Modes need not necessarily correspond with time
- Allows a social network to be defined in time including its duration





APPLICATIONS

- Constructing time dependent routing in mobile opportunistic networks
- Modelling realistic spread of disease
- Social network clustering
- Advantages:
  - Directly tracking communities via the eigen-space
  - Spanning tree construction is easy to implement and may be decentralised