Learning Networks of Heterogeneous Influence

W61

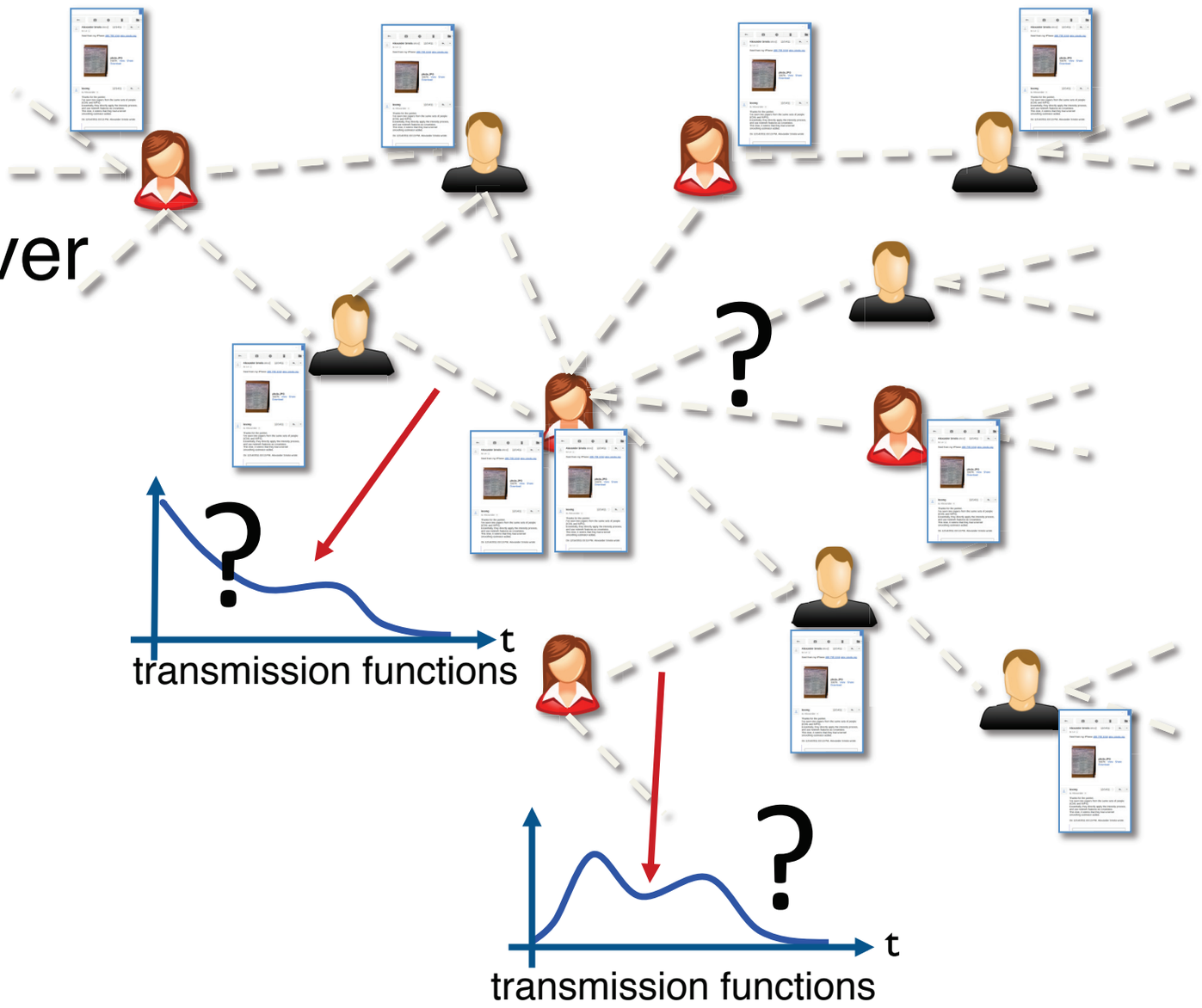
Nan Du[1] (presenter), **Le Song**[1], **Alex Smola**[2], **Ming Yuan**[1]

[1]Georgia Institute of Technology

[2]Google Research

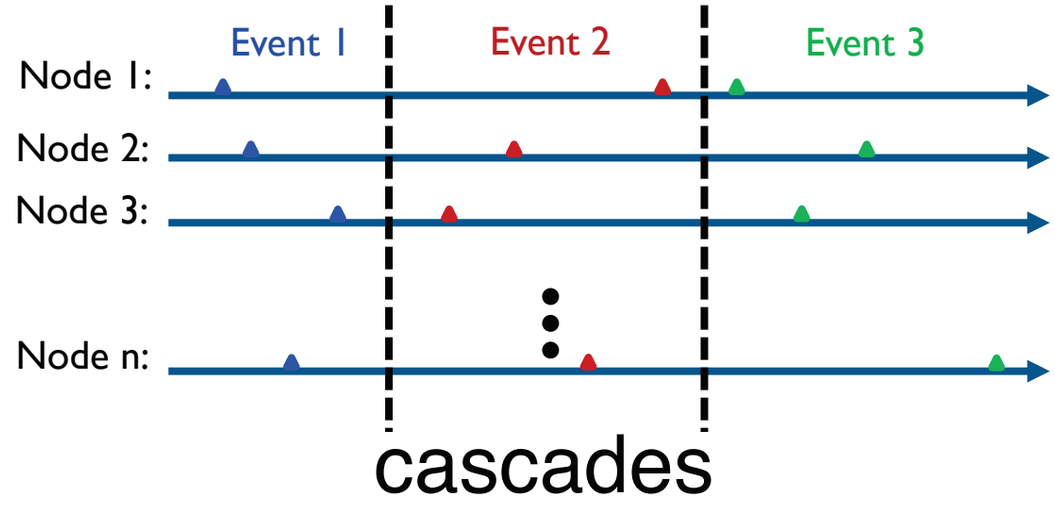
Problem: Diffusions occur over latent networks

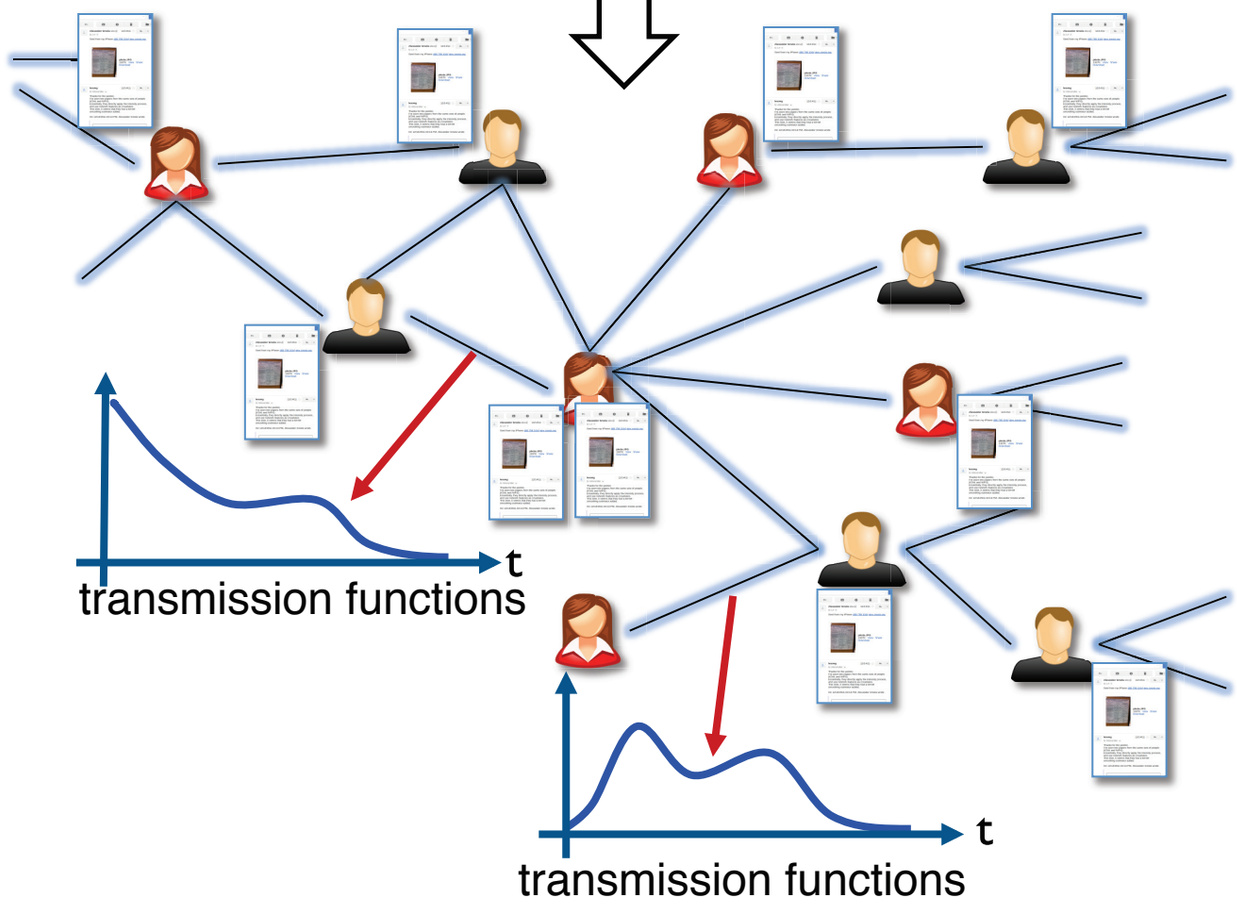
- Only observe infection time
- Unknown sources
- Unknown heterogeneous transmission functions



Goals

- Input: cascades of infection time
- Output:
  - Recovered network structure
  - Inferred transmission functions
- Potential applications:
  - Identify influential individuals
  - Identify important interactions
  - Maximize influence
  - Minimize infections
  - Understand information/virus propagation





Key Ideas

W61

- Use the survival analysis to model diffusion processes from cascade data
- Represent the hazard function h_{ji} between a pair of node by a weighted combination of kernels

$$h_{ji}(\Delta_{ji}) = \sum_{l=1}^{m} \alpha_{ji}^{l} k(\tau_l, \Delta_{ji}),\ \vec{\alpha}_{ji} \geq 0,\ \ \forall j$$

- Estimate the model parameters by group-lasso type of regularization

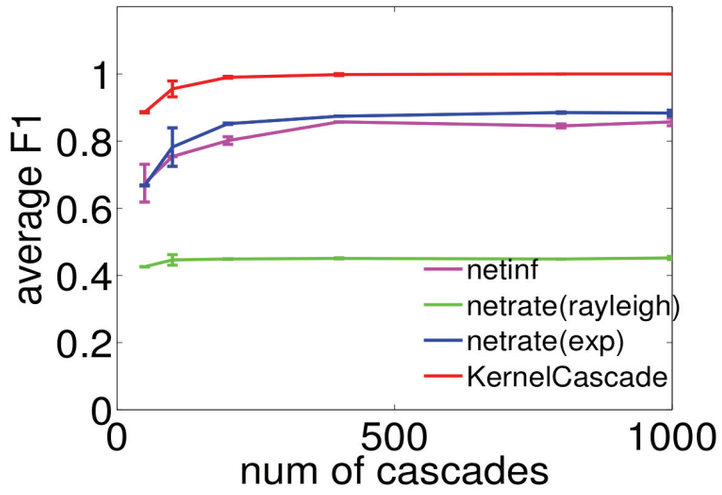
$$\min_{\{\vec{\alpha}_{ji}\}_{j=1}^{N}} \mathcal{L}_i\left(\{\vec{\alpha}_{ji}\}_{j=1}^{N}\right) + \lambda\left(\sum_j \| \vec{\alpha}_{ji} \|\right)^2$$

- Develop an efficient convex optimization algorithm, referred to as **KernelCascade**

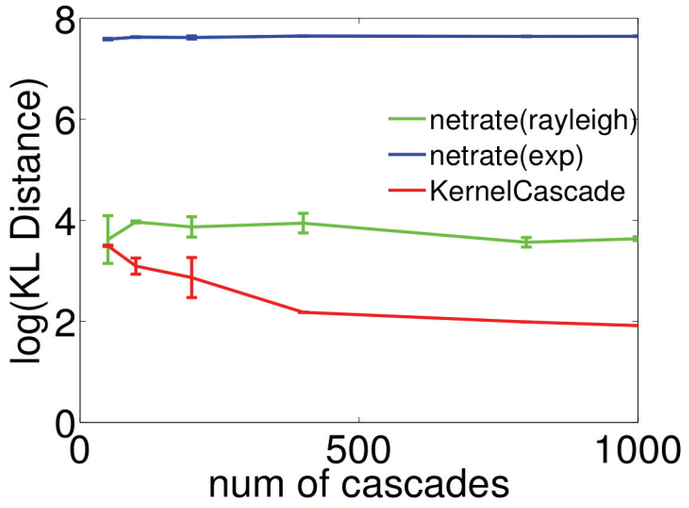
Experimental Results W61

- Significant improvement over previous state-of-the-arts
  - Synthetic datasets (Random, Core-periphery networks)

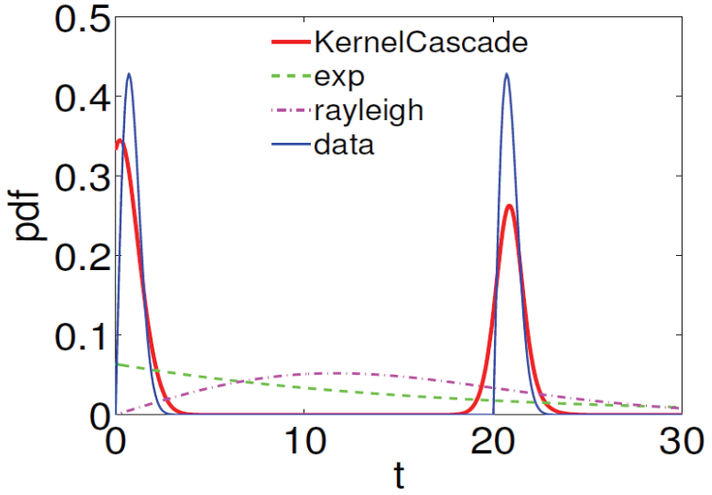
Network recovery



Transmission function recovery

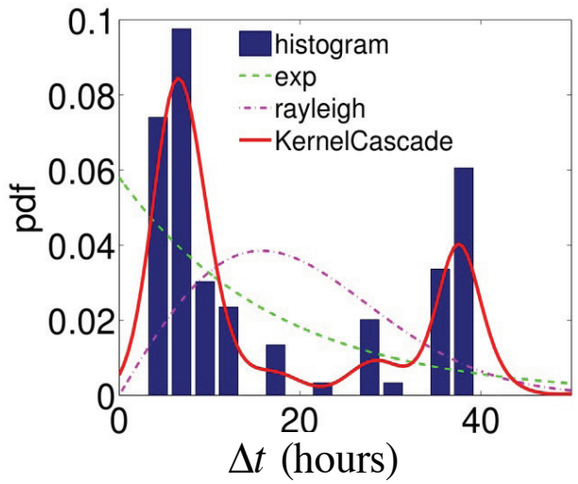


Comparison of recovered functions

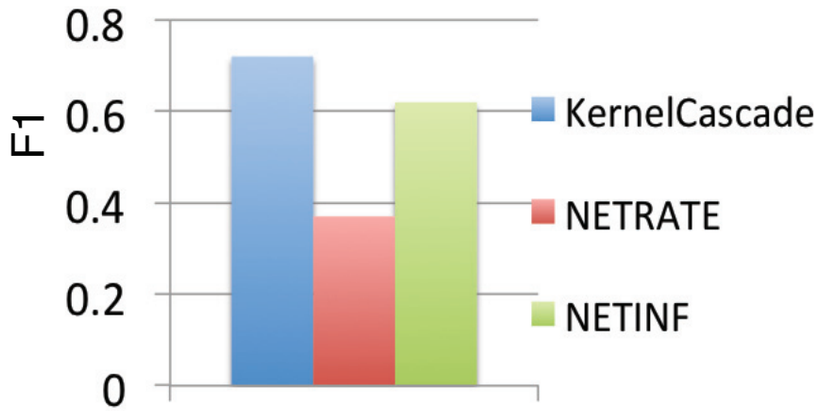


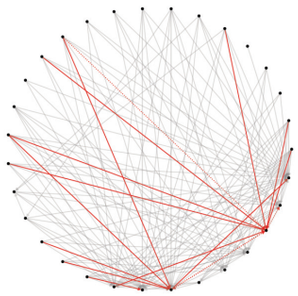
  - Real datasets (Post reference dataset)

Fitted heterogeneous interactions

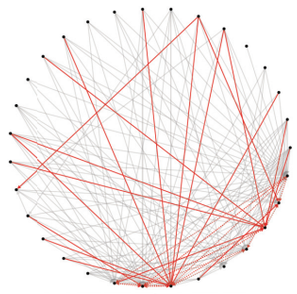


Network recovery

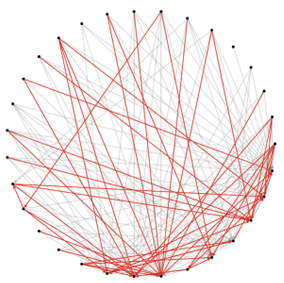




KernelCascade



NETRATE



NETINF