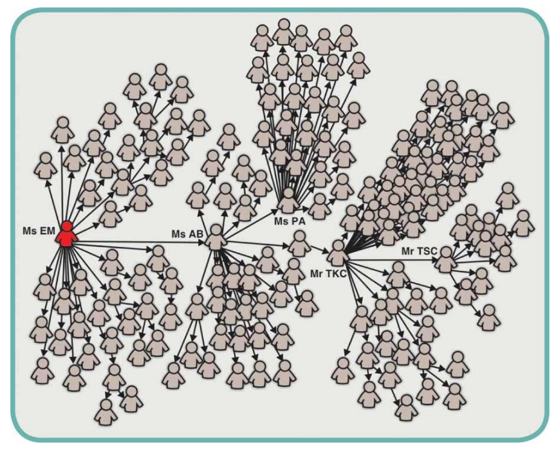
Learning Networks from Dynamics: Detecting Abrupt Changes in Point Processes

Anna Brandenberger, Elchanan Mossel, Ani Sridhar
Massachusetts Institute of Technology

Learning a network from diffusions

Motivating example: network epidemics



Contact network of the first 144 cases of a SARS outbreak in Singapore (src: Normile'13)

- Network structure impacts the spread of epidemics in significant and complex ways
- If the network is known, then we can design effective mitigation measures

- Networks are typically learned through contact tracing ⇒ time-consuming
- Can *key features* of the network be learned in a data-driven manner?

Prior work in learning from dynamics

Exact estimation of networks from diffusions

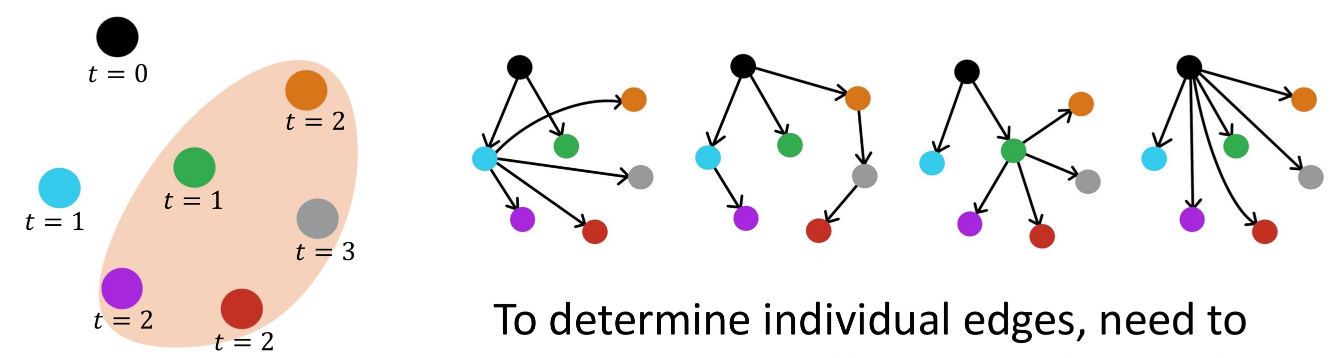
- Early empirical work in epidemiology [Wallinga-Teunis'04] and information flow in blogs [Adar-Adamic'05]
- Scalable and principled methods: the NetInf algorithm [Gomez-Rodriguez et. al.'12]
- Sample complexity of learning networks from cascades
- Convex optimization, message passing, etc.
- Related: learning linear dynamical systems from time series

Common thread: theoretical works are based on optimal estimators (likelihood based) and used to recover the *entire network*

This work: inference of high-degree vertices

Mathematical abstraction: a diffusion (e.g., cascade or epidemic) spreads on an unknown graph. Given the "infection times" for each vertex, learn the network.

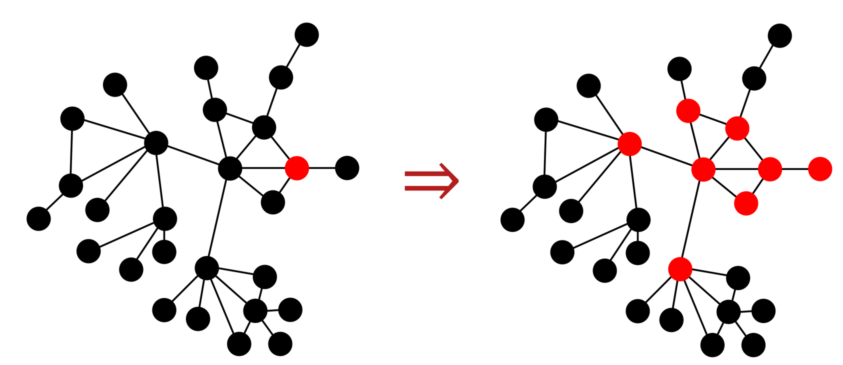
Observations: Many possible explanations:



To determine individual edges, need to observe many diffusions on the graph

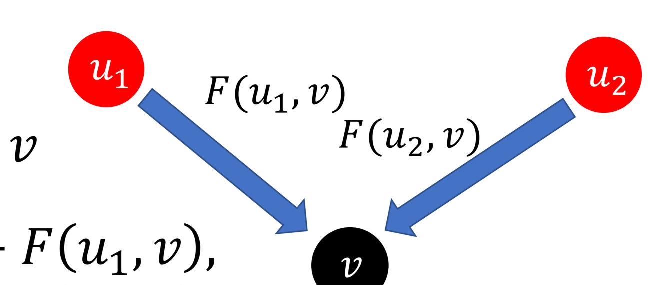
Instead: attempt to infer only high degree vertices
⇒ can this be done in a single diffusion?

The model of continuous time diffusion: Susceptible-Infected model, first passage percolation



λ = (pairwise) spreading rate
$T(v)$ = infection time of vertex v

$$T(v) = \min(T(u_1) + F(u_1, v), T(u_2) + F(u_2, v))$$



Specifically, let $\mathcal{I}(t)$ be the set of infected vertices at time t. Then, for an uninfected vertex v,

$$P\{ v \in \mathcal{I}(t+\epsilon) \mid \mathcal{I}(t) \} = \epsilon\, |\mathcal{N}(v) \cap \mathcal{I}(t)| + o(\epsilon)$$

where $\mathcal{N}(v)$ is the set of neighbors of v.

Model assumptions: given a graph G with n vertices

- At most m (fixed) are *high-degree* (degree $D \geq n^{\alpha}$),
- The rest are low degree (degree $d \leq n^{o(1)}$),
- Two high degree vertices u and v satisfy $dist(u, v) = \omega(1)$ as $n \to \infty$.

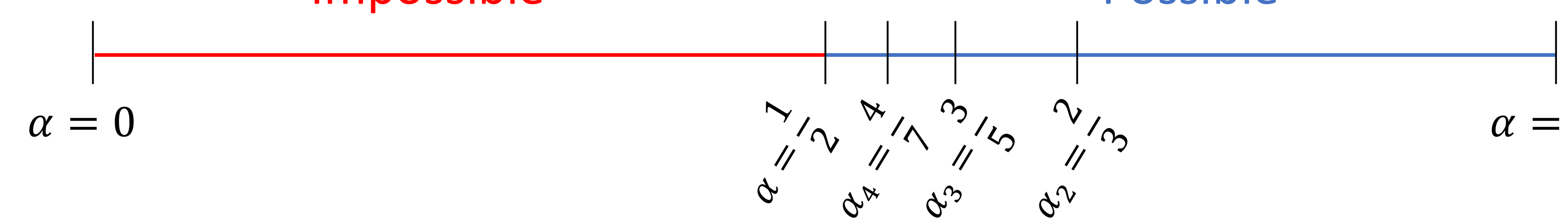
Main result: phase transition in α

Theorem (Mossel-Sridhar'24). Let $\alpha < 1/2$. There exists a distribution μ over admissible graphs such that if $G \sim \mu$ then it is impossible to tell whether there exists a high-degree vertex with probability greater than $o(1)$ as $n \to \infty$.

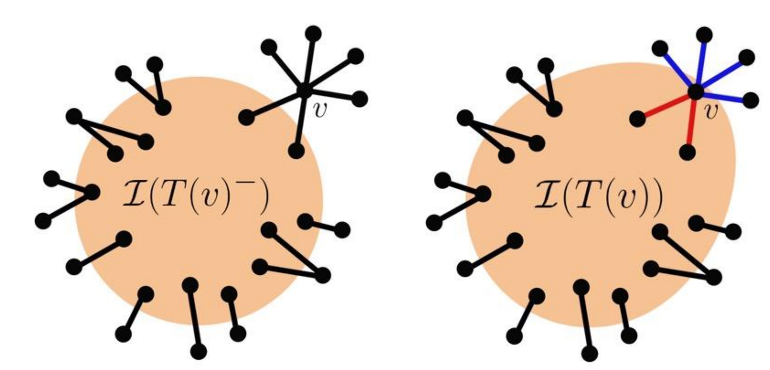
Theorem (B.-Mossel-Sridhar'24+). Let $\alpha > 1/2$. There is an algorithm (depending on α) which outputs a set S of time indices satisfying

$$\{T(v) : \deg(v) \geq D\} \subseteq S \subseteq \bigcup_{v:\deg(v) \geq D} (T(v) - \delta, T(v) + \delta)$$

with probability $1 - o(1)$, where $\delta = 1/poly(n)$.



Main ideas: 1. Second derivative of infection curve $\mathcal{I}(t)$



$$\mathbb{E}\left[\frac{d}{dt} I(t) \middle| \mathcal{F}_t\right] = \lambda \cdot cut(\mathcal{I}(t))$$

$$cut\left(\mathcal{I}(T(v))\right) - cut\left(\mathcal{I}(T(v)^-)\right) = \deg(v) - 2|\mathcal{N}(v) \cap \mathcal{I}(t)|$$

- Second derivative of the infection curve is a (nearly) unbiased estimator for vertex degree. This identifies when a high-degree vertex is infected if $\alpha > 2/3$.

Main ideas: 2. Higher derivatives of infection curve

1. **Local polynomial approximation.** Whp, for every $t \geq 0$ there is a degree $\ell - 1$ polynomial $\bar{I}_{\ell-1}(\cdot, t)$ with $\mathcal{F}_t$-meas. coefficients such that

$$\left|I(s) - \bar{I}_{\ell-1}(s,t)\right| \lesssim n\delta^{\ell} + \sqrt{n\delta}, \qquad s \in (t-\delta, t+\delta),$$

provided no abrupt changes (high-deg infections) in $(t - \delta, t + \delta)$.

2. **Analysis of higher-order derivatives.** We show that

$$\Delta_{\delta}^{(\ell)} I(t) = \boxed{I(t+\delta) - \bar{I}_{\ell-1}(t+\delta, t^-)} \pm O\left(n\delta^{\ell} + \sqrt{n\delta}\right)$$

Residual: Difference between local polynomial prediction based on info before time t and actual process behavior
Note: Residual is similar in magnitude to error terms when no large jumps, by Step 1.

3. **Conditions for detection.** If $\deg(v) \geq D$ then

$$\left|I(T(v)+\delta) - \bar{I}_{\ell-1}(T(v)+\delta, T(v)^-)\right| \gtrsim D\delta$$

The jump can be detected if

$$D \gtrsim n\delta^{\ell-1} + \sqrt{\frac{n}{\delta}} \Rightarrow D \gtrsim n^{\ell/(2\ell-1)} \text{ with } \delta = n^{-1/(2\ell-1)}$$