Immunization Strategies in Complex Networks: Addressing the Problem of Missing Data

Sam Rosenblatt

9/12/2018
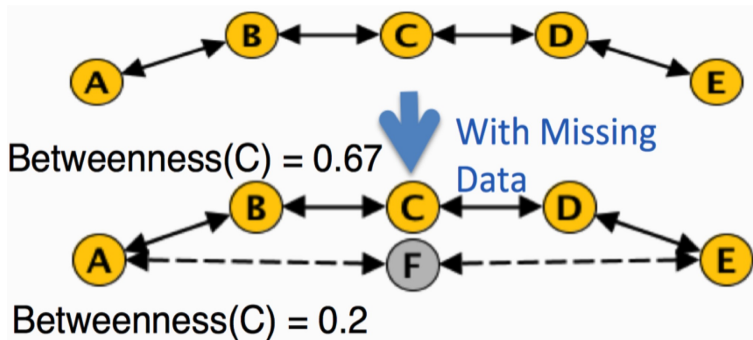
Complete Data is Rare:
Regular Issues with Census Data

- ▶ Unwilling to be interviewed

- ▶ Contact details secret even among partners

- ▶ Population members inaccessible
  - ▶ Incarcerated
  - ▶ Homeless (no address)

Complete Data is Rare:
Hidden Population Example - Colorado Springs

- ▶ At least 44% of nodes missing (out of those found eligible for the study)

- ▶ At least 88% of nodes with incomplete local network data

- ▶ These are generous lower bounds

Complete Data is Rare:
Impact on Centrality Measures

Betweenness(C) = 0.67

With Missing Data

Betweenness(C) = 0.2

Some centrality measures can change drastically with the addition/removal of few nodes

Stochastic Strategies
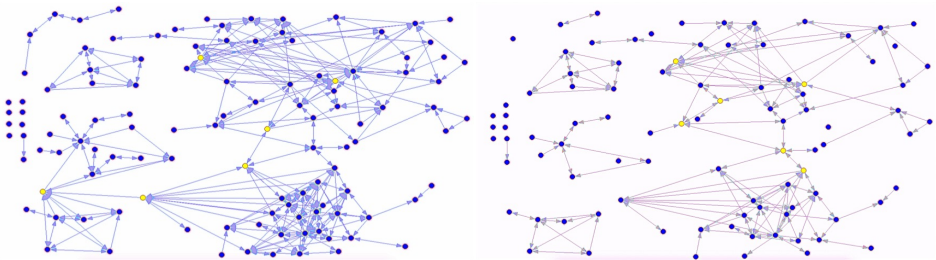
▶ 1: Randomly sample from node list

Stochastic Strategies

- ▶ 1: Randomly sample from node list
- ▶ 2: Investigate local network around randomly sampled node

Stochastic Strategies

- ▶ 1: Randomly sample from node list
- ▶ 2: Investigate local network around randomly sampled node
- ▶ 3: Nodes are found to be central if they are frequently found to be locally important

True Network vs Known Network

Relative Efficacy of Immunization Strategies