Influencer Detection Within Networks

Jacob Becker

Misericordia University

Why Influencer Detection?

- Who are the most “important” nodes in a network?

- Real-world applications:
  - Viral Marketing
  - Epidemiology
  - Online platforms

Agenda

- Background on networks
- Defining influence
- Classic centrality measures
- Structural methods (k-core)
- Influence maximization & modern approaches
- Case study / demo
- Insights & future directions

What is a Network?

- Network is a **graph**
- Graph = **nodes** (vertices) + **edges** (links)

$$G = (V, E)$$

- **Types**: social, biological, technological, or your own.
- **Properties**:
  - Degree = amount of connections
  - Path length = steps between nodes
  - Communities = clusters

Defining “Influence”

- Local Influence = reach neighbors directly
- Global Influence = trigger large-scale cascades

- Not always the same:
  - High-degree node vs “bridge” node

Classic Centrality Measures

- **Degree Centrality**: most connected nodes
- **Betweenness Centrality**: bridges controlling flow
- **Closeness Centrality**: shortest paths to others

Classic Centrality Measures Cont.

- **Eigenvector Centrality (1972)**
  - Importance = connected to other important nodes
  - Recursive definition
- **PageRank (1998)**
  - Probabilistic random-walk version of eigenvector centrality

Structural Methods

- **k-core** = maximal subgraph where every node has ≥ k neighbors
- **Intuition**: nodes in the “core” are structurally embedded
- **Kitsak et al. (2010)**: k-core outperforms degree for identifying spreaders

Influence Maximization

- **Formalized** by Kempe, Kleinberg, & Tardos (2003)
- **Problem**: pick k best seed nodes to maximize spread
- **NP-hard** → greedy algorithms with approximation guarantees

Modern Advances

- Community-aware approaches (multi-scale influence)
- Spectral approaches (eigenvectors, Laplacian, percolation)
- Machine learning models (**GNNs** for influence prediction)
- Dynamic networks (time-varying influence)

Demo

- Comparison of classic measures
- Spread simulation
- See the differences in “influencers” chosen

Results & Insights

(have to do it)

Limitations & Future Directions

- Influence = context-dependent (social vs biological vs tech networks)
- Computational limits for massive networks
- **Open problems**:
  - Temporal dynamics
  - Multilayer networks
  - Interpretable AI-based detection

Conclusion

- Influencer detection = key to understanding complex networks
- Range of methods:
  - Classic
  - Structural
  - Algorithmic
  - Machine Learning
- Takeaway:
  - Must balance **theory**, **scalability**, and **context**

Q&A

- Questions?

beckerj@misericordia.edu

https://github.com/JacobB9990/influencer-detection-workshop.git

References