Spread of Influence in Social Networks
Srinivasan Venkatramanan¹, Anurag Kumar²
Department of ECE, Indian Institute of Science, Bangalore – 560 012
¹vsrini@ece.iisc.ernet.in ²anurag@ece.iisc.ernet.in

Introduction
- Social Network Analysis (SNA) has attracted recent interest
  - due to proliferation of online social networking sites
  - due to availability of large data sets
- SNA also helps in better understanding of
  - Effectiveness of viral marketing
  - Collective human behavior
  - Spread of epidemics

The System Model
- The underlying social network is usually undirected, but the influence graph is assumed to be directed
- Influence matrix \( W = [w_{ij}], \sum w_{ij} \leq 1 \) (usually asymmetric)
- Activation model: Linear Threshold
  - Each node chooses a random threshold \( \theta_j \sim U[0,1] \)
  - Initial set \( A_0 \)
  - \( A_s \) is the terminal set, from which no activation can take place

Existing Work
- [Kempe et al. 2003]
  - The influence maximization problem is NP-hard
  - The influence function is monotonic and submodular in \( A_0 \)
  - Greedy algorithm provides \( 1 - 1/e \) approximate solution
- Cooperative game theoretic solutions (Shapley value), and graph theoretic heuristics (degree, distance) have also been suggested

Recursive expression for \( \sigma(N,A) \)
\[
\sigma(N,i) = 1 + \sum_{j \neq i} w_{i,j} \sigma(N \setminus \{i\}, j)
\]

Examples (Acyclic graphs)
Consider an acyclic graph with \( w_{ij} = 1/\text{degree}(j) \).
We get the following result:
\[
\sigma(N,i) = \text{degree}(i) + 1
\]
Also, one can show that the greedy solution is to pick the nodes in the order of decreasing degree.

DTMC Interpretation
Reversing the edges in the influence graph results in a DTMC

Examples (Complete graphs)

Future Work
- Understand influence spread with other activation models
- Look at other problem variants such as:
  - finding critical links for information spread
  - finding critical nodes for immunization in the context of epidemic spread

References
[1] P. Domingos and M. Richardson, "Mining the Network Value of Customers", SIGKDD’01

Simulation Results
In the above examples, the solutions provided also matches the Noderanking solution inspired by Pagerank.

Data source: Co-authorship network data from NetScience