Modeling the Spread of Influence in Social Networks

Kurtis Taylor - Department of Computer Science
Advisor: Dr. Zesheng Chen

INDIANA UNIVERSITY
PURDUE UNIVERSITY
FORT WAYNE
IPFW

Overview

- Influence can spread through an online social network, which helps spread news, products, and ideas
- Important in understanding scenarios in online social networks, such as the influence maximization problem
- Accurate mathematical model outlines the spread of information and leads to new policies and protocols
- Goal: Derive an accurate model to characterize spread of influence in online social networks
- Challenge: Characterize the spatial dependency between nodes in a social network

Independent Cascade Diffusion Process

- Described by widely popular SIR model (Susceptible → Infected → Recovered)
- Three statuses: Susceptible, Infected, Recovered
- Single chance to infect neighbor *j* with probability $\beta_{ij}(t)$
- Once recovered, node cannot become infected again

Mathematical Framework

- Independent model, used in previous works
- Assumes all nodes are independent, simplest way to approximate distribution

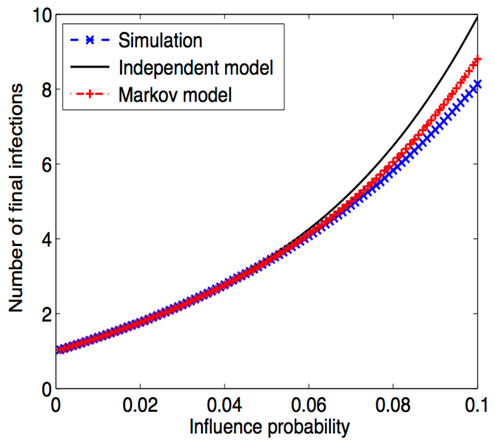
$$P(\mathbf{X}_{N_i}(t) = \mathbf{x}_{N_i}(t) | X_i(t) = 0) = \prod_{j \in N_i} P(X_j(t) = x_j(t))$$

- Proposed Markov model, inspired by the local Markov property
- Given node *i* is susceptible, assumes statuses of its neighbors are conditionally independent

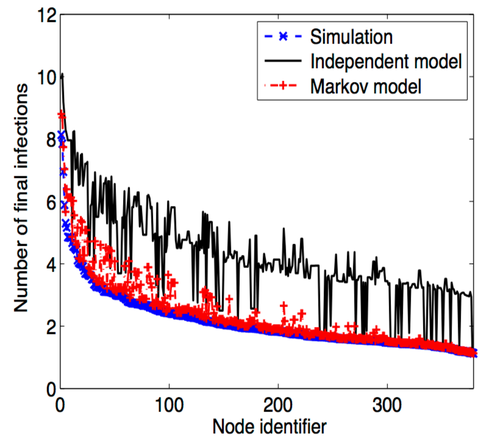
$$P(\mathbf{X}_{N_i}(t) = \mathbf{x}_{N_i}(t) | X_i(t) = 0) = \prod_{j \in N_i} P(X_j(t) = x_j(t) | X_i(t) = 0)$$

Simulation Results

- Coauthorship network of scientists, where author X wrote a joint work with author Y
- Not connected, so only the giant component is considered in our experiment
- 379 total nodes, with each node having an average nodal degree of 4.82



- Initially infected node has largest nodal degree
- Plots how final infections changes with influence probability β



- Every node selected as the initial node when β = 0.1
- Shows the number of final infections

Performance Evaluation of Models

- Independent model is closely related to our proposed Markov model, but there are profound differences in the performance of each model on undirected topologies
- When β is small, both models work well, but not practical for modern online social networks
- As β gets larger (β > 0.06), Markov model more accurately estimates the spread potential of the infection
- Independent model overestimates the spread ability of the infection
- Markov model can accurately predict the number of final infections when the starting node is an arbitrary node

References And Acknowledgements

1. This work is supported in part by IPFW Undergraduate Summer Research Support Program Grant
2. Z. Chen and C. Ji, "Spatial-temporal modeling of malware propagation in networks," *IEEE Transactions on Neural Networks: Special Issue on Adaptive Learning Systems in Communication Networks*, vol. 16, no. 5, Sept. 2005.
3. A. Srivastava, C. Chelmis, and V. K. Prasanna, "Influence in social networks: A unified model?" in *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, Aug. 2014, Beijing, China, pp. 451 - 454.
4. M. Zhang, C. Dai, C. Ding, and E. Chen, "Probabilistic solutions of influence propagation on social networks," in *Proc. of the 22nd ACM international conference on Information & Knowledge Management*, Oct. 2013, San Francisco, CA, pp. 429-438.