Information Diffusion and External Influence in Networks

Seth A. Myers
Stanford University
samyers@stanford.edu

Chenguang Zhu
Stanford University
cgzhu@cs.stanford.edu

Jure Leskovec
Stanford University
jure@cs.stanford.edu

Introduction

- We model the effect of an unobservable external influence on the diffusion of information across a social network.
- Traditional models only consider pieces of information, or contagions, passing from user to user within the network.
- Here, we quantify other sources of information such as mass media that also expose users to contagions.
- We fit our model to URLs diffusing across the Twitter social network.

Modeling Network Diffusion

- Each exposure changes a user's perception of the contagion.
- More exposures validate rumors → more likely to adopt
- More exposures make news stale → less likely to adopt
- The exposure curve [Romero '11] gives probability of adopting the contagion given the number of exposures:
  \[ P(\text{Adoption of contagion after x exposures}) = \frac{\rho_1}{\rho_2} \cdot x \cdot \exp \left( 1 - \frac{x}{\rho_2} \right) \]

Modeling External Influence

- We assume a single external source influencing all users
- The external source's influence is quantified over time, and this is called the event profile:
  \[ \lambda_{\text{ext}}(t) = P(\text{Any user receiving external exposure at time t}) \]

The Complete Model

- For each network neighbor adopting the contagion, the user receives an exposure.
- The external source randomly generates exposures for each user according to the event profile.
- With each exposure, the exposure curve is sampled to decide whether or not the user adopts the contagion.

Fitting the Model

- Given the network and adoption times of each user, the model infers the event profile and exposure curve.
- We iterate between fitting the exposure curve parameters and event profile
- The event profile is fit by matching the expected number of non-adopting nodes to actual number at each time:
  \[ S(t_i) = \sum_{t_{i-1}}^{N} \exp \left[ -\int_{0}^{t_i} \lambda_{\text{ext}}(s) ds \right] \]
  where \( \Lambda_{\text{ext}}(t_i) = \int_{0}^{t_i} \lambda_{\text{ext}}(s) ds \)
  and \( \Lambda_{\text{ext}}(t_i) \) is the number of neighbors of user \( i \) that have adopted the contagion.
- The exposure curve parameters are fit using maximum likelihood of each user’s adoption time.

Real World Data - Twitter

- Across all tweets on Twitter in Jan. 2011 we observed:
  - 18,000 high-volume URLs tweeted a total of 2.67M times by 1.08M users
  - 103M edges connecting users
  - Each URL is a different contagion; we fit the model to each one independently.

Results

- The inferred event profiles capture real world events: