A Probabilistic Approach to Rumor Source Detection and Graph-based Message Passing Algorithms

Liang Zheng and Chee Wei Tan

City University of Hong Kong

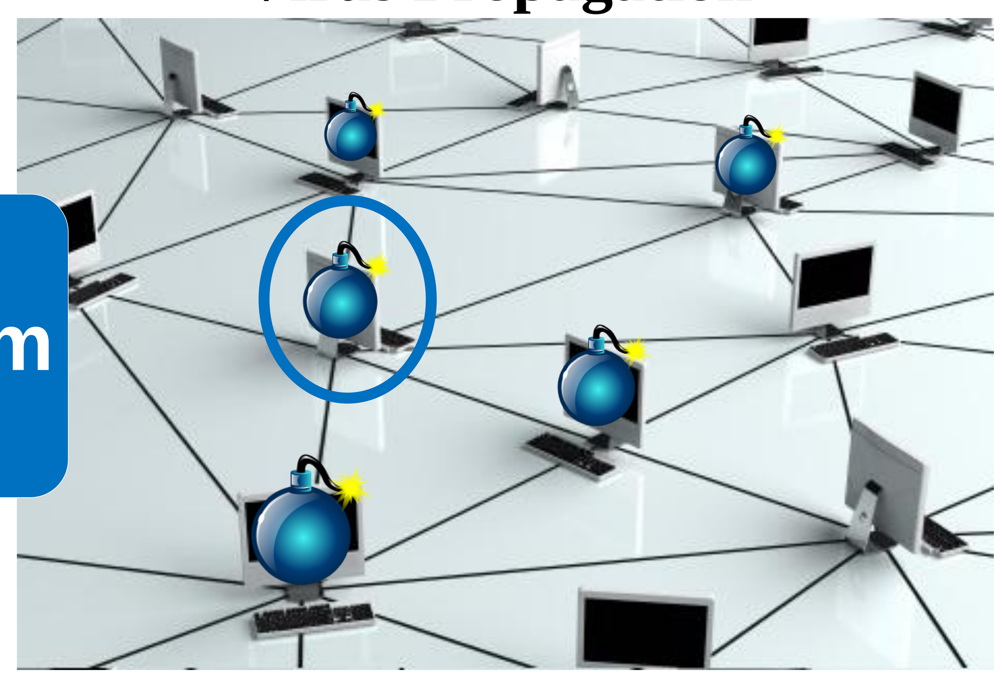
Motivation

Suppose that a rumor originating from a single source spreads in a network. How to identify this rumor source reliably?

Social Network Rumor Spreading

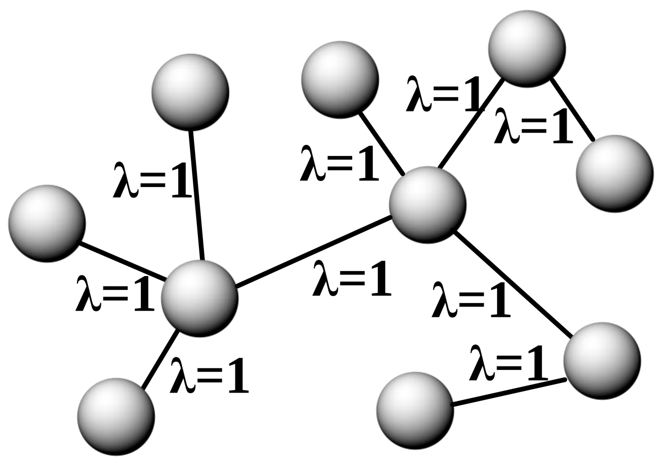
Root out source from Suspects

Computer Network Virus Propagation



System Model

- We consider an infinite network modeled as an undirected graph $G = (V, E)$, where $V = \{s_1, s_2, \dots\}$ is a countably infinite set of nodes and E is the set of edges of the form (i, j) for nodes s_i and s_j in V.
- SI (Susceptible-Infected) spreading model, consisting of (i) susceptible nodes that are capable of being infected; and (ii) infected nodes that can spread the rumor to their immediate neighbors.
- Time to infect a neighbor is exponentially distributed with $\lambda = 1$



A Probabilistic Approach for Tree Graphs

Given the observation time of a rumor graph, the *Maximum Likelihood* (ML) estimator of s^* that maximizes the correct detection probability is given by

$$\hat{s} \in \arg\max_{s \in G_n} P(G_n \,|\, s, T), \tag{1}$$

where $P(G_n \,|\, s, T)$ is the probability of observing G_n at time T supposing that s is the rumor source.

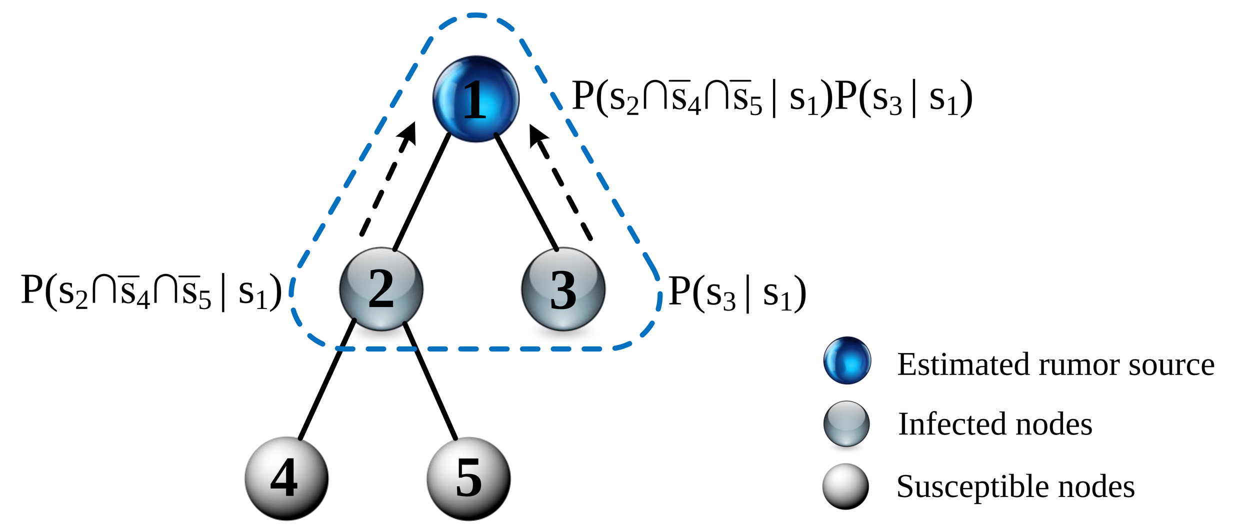
In a tree graph, given source s_j and at time T, the probability of a node s_i being a leaf node in the rumor graph G_n is given by

$$P\left(s_i \bigcap_{s_l \in \text{child}(s_i)}^{L} \bar{s}_l \,\middle|\, s_j\right) = \int_0^T \frac{t^{K_{ij}-1} e^{-t}}{(K_{ij}-1)!} e^{-(T-t)(d_i-1)} \mathrm{d}t, \tag{2}$$

where K_{ij} is the depth of node s_i if s_j is the tree root and d_i is the degree of node s_i. Hence, the ML extimator for a tree can be expressed by

$$\hat{s} \in \arg\max_{s \in G_n} P(G_n \,|\, s, T) = \arg\max_{s \in G_n} \prod_{s_i \in \text{leaf}(G_n|s)}^{L} P\left(s_i \bigcap_{s_l \in \text{child}(s_i)}^{L} \bar{s}_l \,\middle|\, s\right). \tag{3}$$

Graph-based Message Passing Algorithms

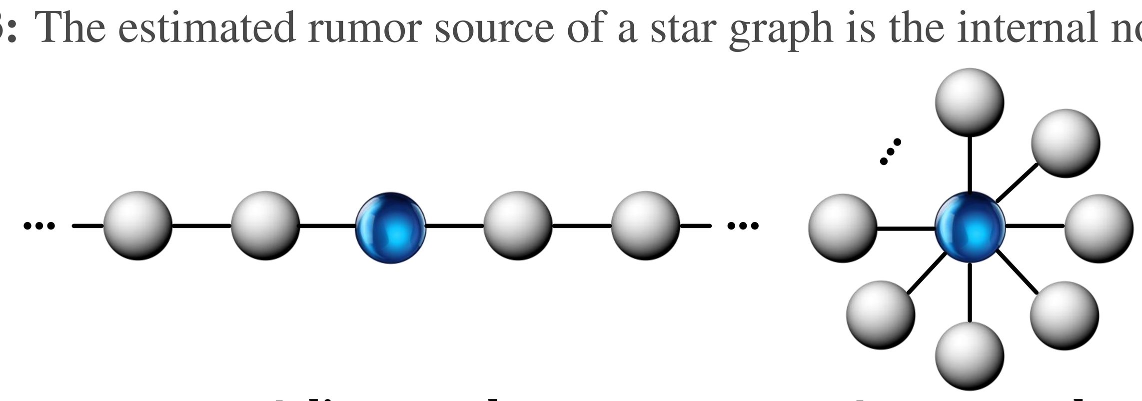


Message-Passing Algorithm

Choose a root node $s \in G_n$.
for $s_i \in G_n$ **do**
  if s_i is a leaf **do**
    Calculate its probability according to (2).
  else if s_i is not a leaf or the root **then**
    Pass the product of the values received from all its child nodes to its parent node.
  else
    The product of the values received from the root's child nodes is the ML.
  end if
end for

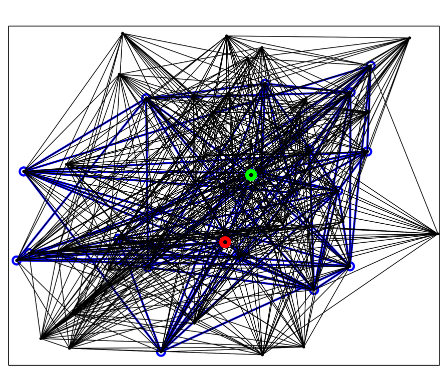
A Property and Special Cases

Proposition 1: Leaf nodes are not the estimated rumor source.
Proposition 2: The estimated rumor source of 2-degree regular tree is the node(s) in the middle of the line.
Proposition 3: The estimated rumor source of a star graph is the internal node.

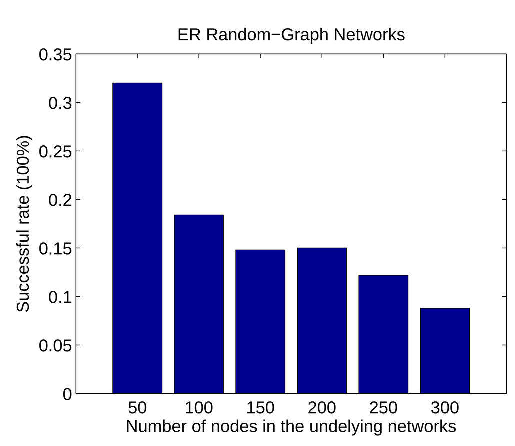


A line graph **A star graph**

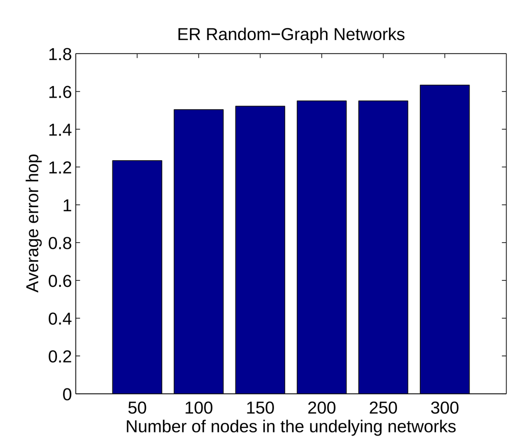
Experiments in Complex Networks



An illustration of rumor spreading in an ER random-graph network.



Correction detection rate of the rumor source detection using (3).



Hops between the actual rumor source and the estimated rumor source.

Related Works

D. Shah and T. Zaman, Rumors in a network: Who's the culprit?, IEEE Transactions on Information Theory, vol. 57, no. 8, pp. 5163-5181, 2011.

L. Zheng, and C. W. Tan, A Probabilistic Characterization of the Rumor Graph Boundary in Rumor Source Detection, IEEE DSP, 2015.

香港城市大學
City University of Hong Kong