Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

A Setting for Rumor Containment Using Linear Threshold Models

Lan Yang * **Ziyue Ma** * **Zhiwu Li** *,** **Alessandro Giua** ***

* *SEME, Xidian University, Xi'an, China*
** *ISE, Macau University of Science and Technology, Taipa, Macau*
*** *DIEE, University of Cagliari, Cagliari, Italy*
(e-mail: yanglan3009@gmail.com, maziyue@xidian.edu.cn, zhwli@xidian.edu.cn, giua@unica.it).

Abstract: Rumor spreads fast in social networks and may produce significant damages to the society. Blocking users in online social networks is normally used as a technical measure to control information spread. In this work, we provide a non-linear formulation to minimize rumor spread in linear threshold networks by blocking a subset of nodes in the network.

Keywords: Social network, Optimization, Rumor containment, Linear Threshold network.

1. INTRODUCTION

The prevalence of online social media has been witnessed in the past decades where people are more and more likely to exchange ideas, share information, and even adopt innovations or new products. Thus a rich body of the research has been devoted to analyzing the propagation of information, opinion, social behavior, and innovation. To model such propagation, Kempe et al. (2003) firstly proposed two main discrete diffusion models, namely the *Linear Threshold Model* (LTM) and the *Independent Cascade Model* (ICM). Both models consist in directed graphs denoted by $G(V, E)$, where each node represents a user or an individual and has two alternative states at a given time step, namely active (if it has adopted the innovation) and inactive (if it has not adopted the innovation). Initially, all the nodes are inactive. At time step $t = 0$, a subset of nodes $S \subseteq V$ is activated in order to start the diffusion process, other nodes stay inactive.

We know that rumors spread very fast in social networks and could produce significant damages to the society. For instance, the rumor "Two explosions in White House and Obama is injured" occurred in April 23, 2013 caused 10 billion USD losses before the rumor was clarified. Therefore, the problem of containing or controlling rumor spread which we focus on in this paper is nonnegligible. Rumor control strategies can be divided into two categories: network disruption strategies and counterbalance strategies.

Counterbalance strategies aim to reduce the diffusion of rumors by spreading correct information. The works in (He et al., 2012; Yang et al., 2020; Zhang et al., 2015) address the problem of minimizing rumor spread by spreading correct information under different extensions of the LTM. In both (He et al., 2012) and (Zhang et al., 2015), it is assumed that their diffusion model is progressive, i.e., an individual activated by a type of information cannot switch to any other ones. However, Yang et al. (2020) allow individuals who activated by rumor first to reconsider their belief which fits better with real individual behaviors.

Network disruption strategies aim to disconnect the inactive nodes from active nodes and can be carried out by removing (or blocking) some critical nodes or links from the underlying network to suppress the rumor spread.

The blocking of a link is understood as deleting the link. Kimura et al. (2009) used the natural greedy algorithm by removing links to find approximate solutions for minimizing the spread of undesirable entities under the ICM and LTM respectively. Yan et al. (2019b) studied the same problem under the ICM. Two different problems to minimize rumor spread are formalized for threshold models in (Kuhlman et al., 2013) by blocking some links with heterogenous costs.

Blocking users in online social networks may refer to denying access to some users such that they cannot see and spread rumor. From the network perspective, the blocking of a node is normally understood as deleting or removing this node and all the connections related with this node from the underlying network such that the rumor cannot pass from it and to other nodes any more. This point of view is considered in works (Wang et al., 2013; Yan et al., 2019a). Wang et al. (2013) use the option of blocking a subset of nodes (so called *blockers*) to minimize rumor spread and a natural greedy approach is presented. In (Yan et al., 2019a), Yan et al. propose a heuristic to find the top-k blockers for general networks and a dynamic programming for tree networks under the framework of the ICM.

We notice that the removal (or block) of nodes involves the removal (or block) of links. However, the solutions to the problem of removing links cannot be directly applied to the problem of removing nodes. Furthermore, in population networks, it is more reasonable and of interest to adopt the option of blocking nodes than blocking links. Therefore, in this work, we adopt the option of blocking nodes to contain rumor spread.

In this paper, we consider the Linear Threshold model to describe the rumor spread in social networks. From the network perspective, in order to identify the set of nodes to be blocked, instead of removing the selected nodes from the network, we increase their thresholds to a value greater than 1. Since, in the LTM it is assumed that the sum of influence weight from one's in-neighbors is 1, any node with threshold that is above 1 will never be activated. We define the top-k blockers problem. Based on the blocked linear threshold network, we present a non-linear programming formulation to find the top-k blockers to minimize the rumor spread.

2. LINEAR THRESHOLD NETWORK

First, we introduce the Linear Threshold model to describe the information propagation in social networks.

A *linear threshold network* N_{LT} is a 4-tuple (V, E, θ, w) where $V = \{1, 2, ..., n\}$ is the set of nodes in the network and $E \subseteq V \times V$ is a set of directed arcs, i.e, $(i, j) \in E$ when there is an arc from node i to another distinct node j. Function $\theta : V \to (0, 1]$ is a mapping that assigns a *threshold value* $\theta_i \in (0, 1]$ to each node $i \in V$. Function $w : V \times V \to (0, 1]$ is a mapping that assigns an influence weight $w_{ij} \in (0, 1]$ to each arc $(i, j) \in E$ such that $w_{ij} = 0$ if $(i, j) \notin E$ and $\sum_{i \in V} w_{ij} = 1$ for all $j \in V$.

Let $\Theta = \text{Diag}([\theta_1, \theta_2, \cdots, \theta_n])$ be the *threshold matrix* whose diagonal elements are the thresholds of the nodes and all other elements are equal to 0. The *weighted adjacency matrix* $W \in [0, 1]^{n \times n}$ of network G is defined as follows,

$$W(i, j) = \begin{cases} w_{ij}, & \text{if } i \neq j \wedge (i, j) \in E \\ 0, & \text{otherwise,} \end{cases}$$

$$W(j, j) = \begin{cases} 1, & \text{if } \sum_{i \neq j} w_{ij} = 0 \\ 0, & \text{otherwise.} \end{cases}$$

Each node in a network represents an individual or agent and the thresholds θ_i represent the different tendencies of nodes to adopt the innovation when their neighbors do (Kempe et al., 2003). We define the *in-neighbor set* of node $i \in V$ as $\mathcal{N}_i = \{j | (j, i) \in E\}$. Arc (i, j) in a network denotes that node j can be influenced by node i.

Let ϕ_0 be the *seed set* which represents a set of agents that are initially activated at step $t = 0$. The activation from the seed set propagates in the network step by step. We denote ϕ_t the set of nodes which are *activated* at step t. The set of nodes *active* at step t, i.e., those that have been activated at step t or at an earlier one, is denoted by $\Phi_t = \bigcup_{k=0}^{t} \phi_k$. By definition, we have $\Phi_0 = \phi_0$.

At each step $t = 1, 2, \ldots$, an inactive node i becomes active if the total influence weight of its neighbors active at step $t - 1$ is at least θ_i, i.e.,

$$i \in \phi_t \Longleftrightarrow \sum_{j \in \mathcal{N}_i \cap \Phi_{t-1}} w_{ji} \geq \theta_i \quad (\forall i \in V \setminus \Phi_{t-1}). \quad (1)$$

The evolution propagates until no more individuals adopt the innovation and the network reaches a steady state. Then we define the set of *final adopters* as $\Phi^*(N_{LT}, \phi_0) = \bigcup_{k=0}^{\infty} \phi_k$.

We can compute the set of final adopters by simulating the evolution process from step 0 to the end at which no more nodes can be activated, i.e., Algorithm 1, with computation complexity $O(nd)$ where d is the average degree of the underlying network.

Algorithm 1 Computing $\Phi^*(N_{LT}, \phi_0)$

Input: A linear threshold network $N_{LT} = (V, E, \theta, w)$, a seed set $\phi_0 \subseteq V$.
Output: The set of final adopters $\Phi^*(N_{LT}, \phi_0)$.
Let $\Phi = \phi_0$, $\Phi_c = V \setminus \phi_0$, $\Phi_{old} = \emptyset$.
Let $k = 0$.
while $\phi_k \neq \emptyset$ **do**
  Let $k = k + 1$, $\phi_k = \emptyset$.
  Let $\Phi_{old} = \Phi$.
  for $v \in \Phi_c$ **do**
    if $\sum_{u \in \Phi_{old} \cap N_v} w_{uv} \geq \theta_v$ **then**
      $\phi_k = \phi_k \cup \{v\}$.
    end if
  end for
  Let $\Phi = \Phi \cup \phi_k$.
  Let $\Phi_c = \Phi_c \setminus \phi_k$.
end while
Let $\Phi^*(N_{LT}, \phi_0) = \Phi$.

2.1 Properties of linear threshold network

Instead of simulating the evolution process of the network to compute the set of final adopters, in this part we show a linear algebraic way based on the notion of cohesiveness.

In (Acemoglu et al., 2011), Acemoglu et al. presented a definition of cohesive set for un-weighted linear threshold networks, we generalize the definition to weighted networks in the following.

Definition 2.1. A subset $X \subseteq V$ is called a *cohesive set* if for all $i \in X$ it holds:

$$\sum_{j \in X \cap \mathcal{N}_i} w_{ji} > 1 - \theta_i \quad (2)$$

Note that for a cohesive set X, if $\phi_0 \cap X = \emptyset$, then $\forall t \geq 0$, $\phi_t \cap X = \emptyset$. In other words, if no individual in X belongs to the seed set, then no individual in X will adopt the innovation at the following steps. The union of cohesive sets is also cohesive.

Lemma 2.1. (Acemoglu et al., 2011) Given a linear threshold network $N_{LT} = (V, E, \theta, w)$ with seed set $\phi_0 \subseteq V$, let $M \subseteq V \setminus \phi_0$ be the *maximal cohesive set* contained in $V \setminus \phi_0$. The final adopter set is:

$$\Phi^*(N_{LT}, \phi_0) = V \setminus M \quad (3)$$

Definition 2.2. Given a set $X \subseteq V$, its *characteristic vector* $\mathbf{x} \in \{0, 1\}^n$ is such that $\mathbf{x}_i = 1$ if $i \in X$, otherwise $\mathbf{x}_i = 0$, i.e.,

$$\mathbf{x}_i = \begin{cases} 1, & \text{if node } i \in X \\ 0, & \text{otherwise.} \end{cases}$$

The sufficient and necessary condition for a cohesive set proposed in (Rosa and Giua, 2013) can also be generalized to weighted networks.

Lemma 2.2. A set $X \subseteq V$ is cohesive if and only if its characteristic vector $\mathbf{x}$ satisfies

$$\mathbf{x}^T W(\cdot, i) \geq 1 - \theta_i \quad (\forall i \in X). \quad (4)$$

and $W(\cdot, i)$ is the i-th column of the weighted adjacency matrix W.

Lemma 2.1 gives a direct way to compute the set of final adopters that does not require to determine the evolution of the network. Then based on Lemma 2.1 and 2.2, a linear characterization of the set of final adopters is proposed shown in Proposition 2.1.

Proposition 2.1. (Rosa and Giua, 2013) Given a linear threshold network $N_{LT} = (V, E, \theta, w)$ with n nodes, let $\phi_0 \subseteq V$ be a seed set with characteristic vector $\mathbf{y}$. The maximal cohesive set M contained in $V \setminus \phi_0$ has a characteristic vector $\mathbf{x}^*$ that is the solution of the following ILP:

$$\begin{aligned} \max_{x} \quad & \mathbf{1}^T \cdot \mathbf{x} \qquad\qquad (ILP-1) \\ \text{s.t.} \quad & \mathbf{1} - \mathbf{x} \geq \mathbf{y} \\ & [I - \Theta - W^T]\mathbf{x} \leq \mathbf{0} \\ & \mathbf{x} \in \{0,1\}^n. \end{aligned}$$

where I is a $n \times n$ identity matrix. The set of final adopters is $\Phi^*(N_{LT}, \phi_0) = \{i \in V | x_i^* = 0\}$.

Note that the operators $\leq$ and $\geq$ are intended componentwise for vectors.

2.2 Network evolution with blocking of nodes

The set of blocked nodes is called as *blocker set* in literature and denoted by $S_b \subseteq V \setminus \phi_0$. We then discuss the effect of blocking of nodes on the network evolution.

Algorithm 2 Computing $\Phi^*(N_{LT}, \phi_0, S_b)$

1: **Input:** A linear threshold network $N_{LT} = (V, E, \theta, w)$, a seed set $\phi_0 \subseteq V$, a blocker set $S_b \subseteq V \setminus \phi_0$.
2: **Output:** The set of final adopters with blocker set S_b: $\Phi^*(N_{LT}, \phi_0, S_b)$.
3: Let $\theta_i^{'} = 0$ for any node $i \in V$.
4: Define a blocked linear threshold network $N_{LT}^{'} = (V, E, \theta^{'}, w)$.
5: **for** $i \in V$ **do**
6:     **if** $i \in S_b$ **then**
7:         $\theta_i^{'} = \theta_i + 1$.
8:     **else**
9:         $\theta_i^{'} = \theta_i$.
10:     **end if**
11: **end for**
12: Compute $\Phi^*(N_{LT}^{'}, \phi_0)$ by Algorithm 1.
13: Let $\Phi^*(N_{LT}, \phi_0, S_b) = \Phi^*(N_{LT}^{'}, \phi_0)$.

Given a linear threshold network $N_{LT} = (V, E, \theta, w)$, a seed set $\phi_0 \subseteq V$, and a blocker set $S_b \subseteq V \setminus \phi_0$, we denote *the set of final adopters with blocker set* S_b by $\Phi^*(N_{LT}, \phi_0, S_b)$ and it can be computed by Algorithm 2. Line 5 – 11 shows the procedure to generate a *blocked linear threshold network* $N_{LT}^{'} = (V, E, \theta^{'}, w)$ where $\theta^{'}$ is a mapping that assigns a threshold value $\theta_i^{'}$ for each node $i \in V$ such that

$$\theta_i^{'} = \begin{cases} \theta_i + 1, & \text{if } i \in S_b \\ \theta_i, & \text{else.} \end{cases}$$

A linear threshold network N_{LT} with blocker set S_b generates a blocked linear threshold network $N_{LT}^{'}$ which is a special linear threshold network with some thresholds larger than 1. Nodes with $\theta_i^{'} > 1$ in $N_{LT}^{'}$ are blocked nodes. The activation rule of a node in $N_{LT}^{'}$ is the same as in the standard LTM N_{LT}. We know from the definition of the LTM that the sum of influence weight from one's in-neighbors is 1. According to Equation (1), nodes with $\theta_i^{'} > 1$ can never be activated, which is consistent with the aim of blocking a node to prevent the node from being activated and in turn influencing its out-neighbors. In the following, we give an example to show how the blocking of nodes influences the network evolution.

We discuss a property of a blocked node in the following which implies that any blocked node is cohesive itself and therefore can never be activated by others.

Proposition 2.2. Given a linear threshold network $N_{LT} = (V, E, \theta, w)$ and a seed set $\phi_0 \subseteq V$, any node in a blocker set $S_b \subseteq V \setminus \phi_0$ is cohesive in network $N_{LT}^{'} = (V, E, \theta^{'}, w)$.

Proof. For any node $i \in S_b$, we have $\theta_i^{'} > 1$. Therefore, it always holds that $\sum_{j \in X \cap \mathcal{N}_i^{in}} w_{j,i} > 1 - \theta_i^{'}$ where X denotes any subset contained in V, which implies that i is cohesive in network $N_{LT}^{'}$. □

3. PROBLEM STATEMENT

Containing rumor spread can be attained by controlling some critical nodes such that the influence of the rumor seeds can be minimized. Here we focus on blocking a set of nodes of the network and consider what we call *top-k blockers problem.*

Problem 3.1. **(Top-k blockers problem)** Given a diffusion model represented by n nodes with seed set $\phi_0 \subseteq V$, let k be a positive integer. Find a set of at most k nodes denoted by $S_b \subseteq V \setminus \phi_0$ to be blocked such that the number of final adopters $|\Phi^*(\phi_0, S_b)|$ is minimized, i.e.,

$$\begin{aligned} \min_{S_b} \quad & |\Phi^*(\phi_0, S_b)| \\ \text{s.t.} \quad & |S_b| \leq k \qquad (a) \\ & S_b \subseteq V \setminus \phi_0 \quad (b) \end{aligned}$$

4. OPTIMAL SOLUTION

Based on the notion of cohesiveness in $N_{LT}^{'}$, we can search for the top-k blockers in linear threshold network by solving a mathematical programming.

4.1 Non-linear formulation

Let $\boldsymbol{b} \in \{0,1\}^n$ be the characteristic vector of S_b where $b_i = 1$ denotes that node $i \in V$ is blocked and 0 not blocked, i.e.,

$$b_i = \begin{cases} 1, & \text{if node } i \text{ is blocked} \\ 0, & \text{otherwise.} \end{cases}$$

Given LTM N_{LT} and a blocker set $S_b \subseteq V \setminus \phi_0$ with characteristic vector $\boldsymbol{b}$, then the threshold of each node $i \in V$ in $N_{LT}^{'}$ can be written as

$$\theta_i^{'} = \theta_i + b_i.$$

Since networks N_{LT} and N'_{LT} have the same set of nodes V and edges E, and influence weights w, we can use the same notations A and W to denote the adjacency matrix and weighted adjacency matrix for N'_{LT}. Then we have the following sufficient and necessary condition for a cohesive set in N'_{LT}.

Corollary 4.1. Given a linear threshold network $N_{LT} = (V, E, \theta, w)$, a seed set ϕ_0, and a blocker set $S_b \subseteq V \setminus \phi_0$ with characteristic vector $\boldsymbol{b}$, a set is cohesive in network $N'_{LT} = (V, E, \theta', w)$ if and only if its characteristic vector $\boldsymbol{x}$ satisfies

$$\boldsymbol{x}^T W(\cdot, i) \geq 1 - (\theta_i + b_i) \quad (\forall i \in X). \tag{5}$$

Let $\Theta' = \text{Diag}([\theta'_1, \theta'_2, \cdots, \theta'_n])$ denote the threshold matrix in network N'_{LT} and can be written as:

$$\Theta' = \Theta + B,$$

where $B = \text{Diag}(\boldsymbol{b})$ is a $n \times n$ matrix (called *blocker matrix*) whose diagonal element $B(i, i)$ equals to b_i and other elements are zero.

Based on Corollary 4.1, we can formalize the top-k blockers problem under linear threshold network as a non-linear programming.

Proposition 4.1. Given a linear threshold network $N_{LT} = (V, E, \theta, w)$ with n nodes, let $\boldsymbol{y}$ be the characteristic vector of the seed set $\phi_0 \subseteq V$ and $k \in \mathbb{R}_+$ a constant. Consider the following non-linear programming (NLP) with binary variable vectors $\boldsymbol{x}$ and $\boldsymbol{b}$:

$$\begin{array}{lll} \max\limits_{\boldsymbol{x},\boldsymbol{b}} & \mathbf{1}^T \cdot \boldsymbol{x} & (NLP) \\ \text{s.t.} & \mathbf{1} - \boldsymbol{x} \geq \boldsymbol{y} & (a) \\ & \boldsymbol{x} \geq \boldsymbol{b} & (b) \\ & \mathbf{1}^T \cdot \boldsymbol{b} \leq k & (c) \\ & [I - (\Theta + B) - W^T] \cdot \boldsymbol{x} \leq \mathbf{0} & (d) \\ & \boldsymbol{x}, \boldsymbol{b} \in \{0, 1\}^n & (f) \end{array}$$

where I is an identity matrix, and let $\boldsymbol{x}^*, \boldsymbol{b}^* \in \{0, 1\}^n$ be the global optimal solution of (NLP). Then the set of nodes $S^*_b = \{i \in V | b^*_i = 1\}$ is the optimal blocker set of the top-k blockers problem and the corresponding set of final adopters is $\Phi^*(N_{LT}, \phi_0, S^*_b) = \Phi^*(N'_{LT}, \phi_0) = \{i \in V | x^*_i = 0\}$.

Proof. Constraint (b) ensures that each blocked node must be also cohesive. Constraints (a) and (b) imply that $\boldsymbol{b} \leq \boldsymbol{x} \leq \mathbf{1} - \boldsymbol{y}$ which indicates that the blocker set S_b and the seed set ϕ_0 are disjoint. Constraint (c) ensures that a set S_b whose characteristic vector is $\boldsymbol{b}$ is an admissible solution to the problem. Constraints (a) and (d) and the objective function ensure that the set M with characteristic vector $\boldsymbol{x}$ is the maximal cohesive set contained in $V \setminus \phi_0$ in $N'_{LT} = (V, E, \theta', w)$, hence $\Phi^*(N'_{LT}, \phi_0)$ has characteristic vector $\mathbf{1} - \boldsymbol{x}$ thanks to Lemma 2.1. Finally, the objective function ensures that the set S^*_b with characteristic vector $\boldsymbol{b}^*$ is an optimal solution to the problem. □

5. CONCLUSION

Blocking users is normally used as a technical measure to control information spread in online social networks. We first present the effect of blocking of nodes on network evolution and then define the top-k blockers problem. To discover the top-k blockers, we provide a non-linear formulation of this problem based on the notion of cohesiveness in linear threshold networks.

REFERENCES

Acemoglu, D., Ozdaglar, A., and Yildiz, E. (2011). Diffusion of innovations in social networks. In *50th IEEE Conference on Decision and Control and European Control Conference*, 2329–2334. IEEE.

He, X., Song, G., Chen, W., and Jiang, Q. (2012). Influence blocking maximization in social networks under the competitive linear threshold model. In *Proceedings of the 2012 SIAM International Conference on Data Mining*, 463–474. SIAM.

Kempe, D., Kleinberg, J., and Tardos, E. (2003). Maximizing the spread of influence through a social network. In *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 137–146. ACM.

Kimura, M., Saito, K., and Motoda, H. (2009). Blocking links to minimize contamination spread in a social network. *ACM Transactions on Knowledge Discovery from Data*, 3(2), 9.

Kuhlman, C.J., Tuli, G., Swarup, S., Marathe, M.V., and Ravi, S. (2013). Blocking simple and complex contagion by edge removal. In *Proceedings of the 13th IEEE International Conference on Data Mining*, 399–408. IEEE.

Rosa, D. and Giua, A. (2013). On the spread of innovation in social networks. *4th IFAC Workshop on Distributed Estimation and Control in Networked Systems*, 46(27), 322–327.

Wang, S., Zhao, X., Chen, Y., Li, Z., Zhang, K., and Xia, J. (2013). Negative influence minimizing by blocking nodes in social networks. In *Proceedings of Workshops at the 27th AAAI Conference on Artificial Intelligence*, 134–136.

Yan, R., Li, D., Wu, W., Du, D.Z., and Wang, Y. (2019a). Minimizing influence of rumors by blockers on social networks: Algorithms and analysis. *IEEE Transactions on Network Science and Engineering*, early access.

Yan, R., Li, Y., Wu, W., Li, D., and Wang, Y. (2019b). Rumor blocking through online link deletion on social networks. *ACM Transactions on Knowledge Discovery from Data*, 13(2), 16.

Yang, L., Li, Z., and Giua, A. (2020). Containment of rumor spread in complex social networks. *Information Sciences*, 113–130.

Zhang, H., Zhang, H., Li, X., and Thai, M.T. (2015). Limiting the spread of misinformation while effectively raising awareness in social networks. In *Proceedings of the International Conference on Computational Social Networks*, 35–47. Beijing, China.