2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

Fight under Uncertainty: Restraining Misinformation and Pushing out the Truth

Huiling Zhang, Alan Kuhnle, *Student Member, IEEE,* J David Smith and My T. Thai, *Senior Member, IEEE*
Department of Computer and Information Science and Engineering
University of Florida, Gainesville, Florida 32611
Email: {huilingzhang, kuhnle, emallson}@ufl.edu, mythai@cise.ufl.edu

***Abstract*—While online social networks (OSNs) have become an important platform for information exchange, the abuse of OSNs to spread misinformation has become a significant threat to our society. To restrain the propagation of misinformation in its early stages, we study the *Distance-constrained Misinformation Combat under Uncertainty* problem, which aims to both reduce the spread of misinformation and enhance the spread of correct information within a given propagation distance. The problem formulation considers the competitive diffusion of misinformation and correct information. It also accounts for the uncertainty in identifying initial misinformation adopters. For competitive propagation with *major-threshold activation*, we propose a solution based on stochastic programming and provide an upper-bound in the presence of uncertainty. We propose an efficient Combat Seed Selection algorithm to tackle *general-threshold activation*, in which we define a measure, *"effectiveness"*, to evaluate the contribution of nodes to the fight against misinformation. Through extensive experiments, we validate that our algorithm outputs high-quality solution with very fast computation.**

I. Introduction

Online social networks (OSNs) are no doubt powerful tools to share information and opinions. However, OSNs have also facilitated the spread of inaccurate and false information. Misinformation covers a wide variety of topics, ranging from political conspiracy theories and deceptive advertising to false reporting during breaking events [1]–[4]. The World Economic Forum has listed massive misinformation as one of the main threats to our society [5]. The wide spread of misinformation can cause irreparable damage to a business's reputation, complicate disease control, and lead to public panic and economic loss. It is therefore crucial to counteract the spread of misinformation in OSNs.

The fight against misinformation is challenging due to the large size of OSNs and intricate connections between users. A growing amount of work has been devoted to misinformation containment. Some has studied misinformation identification or locating misinformation sources [6]–[8]. Meanwhile, others have looked to limit or block misinformation. Basaras et al. proposed solutions to block the propagation of malicious items by removing key connections [9]. Fan et al. presented algorithms to select initial protectors so as to limit the bad influence of rumors. However, both approaches seem likely cause side effects. For instance, blocking users' access on OSNs may lead to confusion and accusations of censorship, while marking massages or articles as misinformation can instead further entrench false beliefs [4], [10].

IEEE/ACM ASONAM 2018, August 28-31, 2018, Barcelona, Spain
978-1-5386-6051-5/18/$31.00

In this paper, we combat the spread of misinformation via a different approach, i.e., disseminating correct information. It has been reported that once false beliefs have been adopted, corrections are rarely adopted since people tend to accept only information that is consistent with their existing knowledge and beliefs [4]. However, mis-perceptions of posts containing misinformation can be significantly reduced by showing related correct information alongside them [10]. Therefore, besides limiting the spread of misinformation, we want to inform users of accurate information and provide them prior knowledge that enables the rejection of misinformation.

Since misinformation can quickly spread throughout OSNs, intervention strategies need to be adopted as soon as possible. We therefore consider a *Distance-constrained Misinformation Combat under Uncertainty* (d-MCU) problem. It aims to disseminate correct information from some selected users so as to restrain the influence of misinformation and promote the adoption of correct information within a given number of propagation hops. The d-MCU problem associates a probability with each user to indicate the likelihood of being the initiator of misinformation. This formulation incorporates the uncertain factors in identifying early adopters of misinformation. For example, the number of network monitors in large-scale OSNs that we are able to deploy is limited, and thus collected data is usually imperfect and incomplete [11], [12]. Moreover, messages can evolve as they circulate on OSNs, which complicates verification. Worse yet, malicious individuals may collude to disseminate misinformation [6].

Beyond the uncertainty associated with initial adopters of misinformation, the challenges in solving d-MCU problem stem from the following aspects. First, different people may not react in the same way to the same piece of information. Second, the diffusion of misinformation and correct information are intrinsically intertwined as they essentially compete for peoples' beliefs. Hence the timing of receiving information is important in shaping users' beliefs. Furthermore, sources of correct information may not be able to reach users under the threat of misinformation. Hence we need to find out a trade-off between blocking misinformation and promoting correct information. In this paper, we first focus on a special scenario, referred as *majority-threshold activation*, where users' states depend on the majority of their friends. Then we address general scenarios where users' acceptance behaviors are heterogeneous. Our main contributions are summarized in the following.

- We define a measure called the *"combat score"* to evaluate the quality of selected initiators of correct information in the presence of uncertainty and show the #P hardness of the computation of combat score.
- For majority-threshold activation, we formulate a two-stage stochastic program, which incorporates uncertainty in identifying misinformation adopters. We obtain an upper-bound by solving this program.
- We propose an efficient algorithm for misinformation containment in general cases. The algorithm keeps track of the measures related to competitive diffusion and produces high-quality solutions.
- Extensive experiments on real-world topologies demonstrate that our proposed algorithm outperforms other heuristics, in terms of maximizing combat score.

Paper Organization. We introduce the model describing the competitive diffusion of misinformation and correct information in Section II. Problem formulation and complexity analysis are given in Section III. Section V focuses on majority-threshold activation and provides an upper-bound. Section VI proposes a scalable solution for general-threshold activation. In Section VII, we conduct experimental evaluation of algorithm performance. We briefly review related works in Section VIII, and conclude the paper in Section IX.

II. Competitive diffusion model

Threshold models are widely used in the literature to describe influence propagation in OSNs [13]–[16]. In this paper, we use a separated-threshold model, which was originally proposed to model product competition [17]. In the rest of the paper, "truth" and "correct information" are used interchangeability. We use A and B to denote the *truth* and *misinformation*, respectively. Consider an OSN represented as a directed graph $G = (V, E)$, where node set V represents users and edge set E represents social interactions among users. Let $n = |V|$ and $m = |E|$. Each edge $(u, v) \in E$ is assigned with two weights, w^A_{uv} and w^B_{uv}, which characterize the impact of node u on node v with respect to the truth and misinformation, respectively. Every node v has two thresholds, θ^A_v and θ^B_v, for the acceptance of the truth and misinformation, respectively.

In this model, each node can be in one of three states, i.e., *inactive*, *truth active (A-active)*, and *misinformation active (B-active)*. A node is A-active if it is exposed to the influence of the truth and really believe the truth. In this paper, we focus on diffusion process unfolding in discrete steps and will consider continuous-time diffusion in future work. Starting from truth initiators I^A and misinformation initiators I^B at $t = 0$, the diffusion process unfolds and gets more and more users activated. We denote the set of A-active nodes at time-step $t > 0$ as S^A_t. Inactive nodes become active if the acceptance threshold is satisfied. Specifically, if $\sum_{u \in N^{in}_v \cap S^A_t} w^A_{uv} \geq \theta^A$, node v will become *A-active* in time-step $t+1$, where N^{in}_v is the set of v's incoming neighbors. B-active and B-activation process are similarly defined.

In the scenario where users are exposed to both correct information and misinformation and both thresholds are satisfied, we assume that users will accept correct information. This assumption incorporates the fact that, in the combat of misinformation, correct information is usually disseminated by authorized parties and can be easily verified. The propagation is progressive, that is, none of active nodes can change back to inactive state. Node states cannot switch between *A-active* and *B-active*, which reflects *continued influence* of information in memory [4]. The process proceeds till no more activation happens. In the diffusion model, we define deterministic thresholds. It is an extension from deterministic linear threshold model, which is harder than its probabilistic counterpart and with less prior work [18], [19].

III. Problem Formulation and Complexity

A. Uncertainty in identifying initial misinformation adopters

In most cases, we do not have complete knowledge about misinformation initial adopters. For instance, the surveillance data on misinformation propagation is sampled from large networks. The imperfect data brings about the uncertainty in determining adopters of misinformation. To accommodate this uncertainty, we use a set of probabilities $\mathbb{P}$ over nodes to describe the partial knowledge about node states. Specifically, each node v is an initial adopter of misinformation with probability p_v, and thus the probability that $I^B \subseteq V$ is the seed set of misinformation is

$$Pr[I^B] = \prod_{a \in V \setminus I^B} (1 - p_a) \prod_{b \in I^B} p_b. \tag{1}$$

If available prior knowledge can help exclude some nodes from initial misinformation adopters, we can just set the probability of these nodes to be 0. When $p_v = 0, \forall v \in V$, the diffusion model reduces to deterministic linear threshold model of single diffusion [18].

B. Problem definition

Once people believe misinformation, they become resistant to correction. As suggested in [4], giving people prior accurate information can help them reject the misinformation. To capture this phenomenon, we consider the combat of misinformation from two aspects, that is, informing users with correct information and saving users who are expected to adopt misinformation without any intervention. Existing literature about competitive diffusion focuses on either the reduction of bad campaigns or the promotion of good campaigns [15], [20], [21]. Therefore, we need to define a new objective function in order to incorporate both aspects. Meanwhile, real applications often require misinformation to be countered before given deadlines. Information diffusion bounded by d propagation-hops is therefore of interest. When propagation times among users in OSNs are homogeneous, the constraint on d can be seen as the time constraint. Our objective function, named *combat score*, is defined as follows:

$$\sigma(I^A, d, \mathbb{P}) = \sigma^A(I^A, d, \mathbb{P}) + (\sigma^B(\emptyset, d, \mathbb{P}) - \sigma^B(I^A, d, \mathbb{P})),$$

where $\sigma^A(I^A, d, \mathbb{P})$ and $\sigma^B(I^A, d, \mathbb{P})$ respectively denote the expected number of *A-active* nodes and *B-active* nodes within d hops under probability distribution $\mathbb{P}$ with respect to truth initiators I^A. $\sigma^B(\emptyset, d, \mathbb{P})$ denotes the expected number of nodes finally adopting misinformation if there is no intervention. Considering all possible misinformation initiator set I^B, we have $\sigma(I^A, d, \mathbb{P}) = \sigma^B(\emptyset, d, \mathbb{P}) + \sum_{I^B} Pr[I^B](\sigma^A_{I^B}(I^A, d) - \sigma^B_{I^B}(I^A, d))$, where $\sigma^A_{I^B}(\cdot)$ and $\sigma^B_{I^B}(\cdot)$ are values corresponding to a realization I^B from $\mathbb{P}$. With the objective of countering misinformation propagation, we define the following problem.

Definition 1. ***Distance-constrained Misinformation Combat under Uncertainty (d-MCU)*** *Given a social network $G = (V, E)$, a probability distribution $\mathbb{P}$ of the misinformation initial adopters, and a budget k, d-MCU problem aims to find a seed set I^A with $|I^A| \leq k$ for the truth such that the combat score $\sigma(I^A, \mathbb{P}, d)$ is maximized.*

Since $\sigma^B(\emptyset, d, \mathbb{P})$ can be viewed as a constant number, in the rest of the paper, we will simplify the combat score as

$$\sigma(I^A, d, \mathbb{P}) = \sigma^A(I^A, d, \mathbb{P}) - \sigma^B(I^A, d, \mathbb{P}).$$

When $\mathbb{P}$, d are known from the context, we can simply denote $\sigma(I^A, \mathbb{P}, d)$ as $\sigma(I^A)$. Similarly, we denote $\sigma^A(I^A, d, \mathbb{P})$ and $\sigma^B(I^A, d, \mathbb{P})$ as $\sigma^A(I^A)$ and $\sigma^B(I^A)$, respectively.

IV. Complexity of Computing Combat Score

Computation of combat score is essential in assessing the quality of selected seeds I^A. If I^A and I^B are given, one can easily compute the value of combat score as the propagation process runs deterministically. In this paper, we focus on the propagation of information within a bounded number of hops, which is different from the typical influence maximization problem [13], [14]. We are thus interested in the complexity of the computation of $\sigma(\cdot)$ when the number of propagation hops d is given. In the following, we show its #P-hardness even for small d, which prohibits exact evaluation of $\sigma(\cdot)$ in polynomial time.

Theorem 1. *The computation of $\sigma(I^A)$ in d-MCU is #P-hard under competitive diffusion model when $d \geq 3$.*

Proof. Given that counting the number of minimal set cover (SC) is a #P-complete problem, we will prove the #P-hardness of computing $\sigma(I^A)$ by a reduction from SC problem.

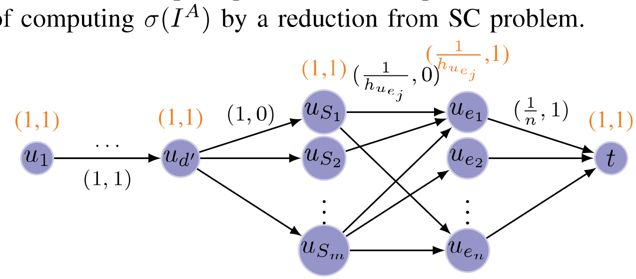


Fig. 1: Reduction from SC. Each node u_{S_i} represents a set S_i and u_{e_j} corresponds to an element $e_j \in U$.

Consider an arbitrary instance of SC (U, S), where U is the ground set of elements and $S = \{S_1, S_2, \cdots, S_m\}$ is a collection of subsets of elements in U. The reduction from SC to d-MCU is defined as follows. As Fig. 1 shows, for each set $S_i \in S$, a node u_{S_i} with thresholds $(1, 1)$ is added into the graph and $V_S = \{u_{S_1}, u_{S_2}, u_{S_m}\}$. For each element $e_j \in U$, a node u_{e_j} with threshold $(1/h(u_{e_j}), 1)$ is added, where $h(u_{e_j})$ is the number of sets having u_{e_j} in it and $V_U = \{u_{e_1}, u_{e_2}, \cdots, u_{e_n}\}$. Then we connect set nodes V_S with element nodes V_U according to their relationship indicated by SC instance itself.

Later we add $d' = d-2$ nodes, denoting as $u_1, u_2, \cdots, u_{d'}$, into G and connect node $u_{d'}$ with all nodes in V_S with weights $(1, 0)$. And we add a node t with thresholds $(1, 1)$ which is connected by nodes in V_U with weights $(1/n, 1)$. Let $I^A = \{u_1\}$ and each node in V_S be a misinformation adopter with probability $p = 1/2$. Suppose there exists an algorithm $\mathbb{A}$ to compute combat score. We can compute the value of $\sigma(I^A)$, denoted as σ_1. Then the combat score without node t is computed, denoted as σ_2. Then the number of set cover can be obtained as $C = 2^{|V_S|}|\sigma_2 - \sigma_1|$. ☐

V. Majority-threshold Activation

In this section, we consider d-MCU with *majority-threshold activation*, i.e., every inactive node gets activated only if the sum of influence from active incoming neighbors is no less than *half* of the total weight from all incoming neighbors. Majority-threshold activation is a generalization of classical majority threshold model [18], [22]. In the following subsections, we propose a solution addressing uncertainty in identifying initial adopters based on stochastic programming.

A. Stochastic Programming

In majority-threshold activation, $\theta^A_v = \frac{1}{2}\sum_{u \in N^{in}_v} w^A_{uv}$ and $\theta^B_v = \frac{1}{2}\sum_{u \in N^{in}_v} w^B_{uv}$. According to the diffusion model, the state of an inactive node only depends on its incoming neighbors. We thus normalize all edge weights of $G = (V, E)$. Specifically, we assign weights $\bar{w}^A_{uv}$ and $\bar{w}^B_{uv}$ to edge (u, v), where $\bar{w}^A_{uv} = \frac{w^A_{uv}}{\theta^A_v}$ and $\bar{w}^B_{uv} = \frac{w^B_{uv}}{\theta^B_v}$. Then all thresholds are set to be 1. We continue using the notation w^A_{uv}, w^B_{uv} on normalized graph for the sake of simplicity.

As there exists uncertainty in determining initial adopters of misinformation, we propose a solution based on two-stage stochastic programming. For each node $v \in V$, we use variable s_v to indicate whether v is selected as a seed of correct information. When v is selected, $s_v = 1$; otherwise, $s_v = 0$. Since budget is limited, we impose on s_v the constraint $\sum_{v \in V} s_v \leq k$. We decide the values of s_v in the first stage.

To track nodes' states when misinformation and correct information propagate simultaneously in the network, we introduce more variables to describe nodes' states as follows. For every $i = 1, \ldots, d$,

$$x^i_v = \begin{cases} 1, & \text{if } v \text{ is } A\text{-active in round } i; \\ 0, & \text{otherwise,} \end{cases} \quad \textit{and}$$

$$y^i_v = \begin{cases} 1, & \text{if } v \text{ is } B\text{-active in round } i; \\ 0, & \text{otherwise.} \end{cases}$$

Due to the lack of full knowledge about misinformation early adopters, y_v^1 ($v \in V$) is a random Bernouli variable and satisfies that $Pr[y_v^1 = 1] = p_v$ and $Pr[y_v^1 = 0] = 1 - p_v$. In each possible realization of misinformation early adopters, the values of $\{y_v^1\}$ are revealed to be either 0 or 1. In the second stage, the influence of correct information can be computed after k seeds are selected. Note that, values of x_v^i ($i = 2, \dots, d$) and y_v^i ($i = 2, \dots, d$) are determined by the diffusion process. Then, we can have the following second stage integer program, denoted as $\sigma(s, x, y)$.

$$\max \sum_{v\in V} x_v^d - \sum_{v \in V} y_v^d \tag{2}$$

$$\text{s.t. } x_v^1 \le s_v, \qquad \forall v \in V \tag{3}$$

$$x_v^1 + y_v^1 \le 1, \qquad \forall v \in V \tag{4}$$

$$x_v^i \ge x_v^{i-1}, \qquad \forall v \in V, i = 2, \dots, d \tag{5}$$

$$y_v^i \ge y_v^{i-1}, \qquad \forall v \in V, i = 2, \dots, d \tag{6}$$

$$x_v^i + y_v^i \le 1, \qquad \forall v \in V, i = 2, \dots, d \tag{7}$$

$$\sum_{u \in N_v^{in}} w_{uv}^A x_u^{i-1} + x_v^{i-1} \ge x_v^i, \qquad \forall v \in V, i = 2, \dots, d \tag{8}$$

$$\sum_{u \in N_v^{in}} w_{uv}^B y_u^{i-1} + y_v^{i-1} \ge y_v^i, \qquad \forall v \in V, i = 2, \dots, d \tag{9}$$

$$\sum_{u \in N_v^{in}} w_{uv}^B y_u^{i-1} - 1 \le y_v^i + x_v^i, \qquad \forall v \in V, i = 2, \dots, d \tag{10}$$

$$s_v, x_v^i, y_v^i \in \{0, 1\} \tag{11}$$

Constraints (4)-(7) enforce the progressiveness and non-compatible acceptance of the truth and misinformation. Constraints (8)-(10) describe the threshold-based activation.

The two-stage stochastic program for the d-MCU problem can be formulated as follows.

$$(M_1) \qquad \max \mathbb{E}_{\mathbb{P}}[\sigma(s, x, y)] \tag{12}$$

$$\text{s.t.} \sum\nolimits_{v \in V} s_v \le k, \tag{13}$$

$$\sigma(s, x, y) \text{ is given in (2)-(11)},$$

B. Solution to Two-stage Program

We can write the program M_1 in a standard form:

$$\max a^T x - c^T y \text{ subject to } Qs + Tx + Wy \le b,$$

where Q, T, W correspond to parameters in M_1. There are $N = 2^{|V|}$ possible realizations of random vector y^1, denoted as $y^{1,1}, y^{1,2}, \dots, y^{1,N}$ with probabilities $\gamma_1, \gamma_2, \dots, \gamma_N$. Since the value of variables x_v^i ($i = 1, \dots, d$) and y_v^i ($i = 2, \dots, d$) can be determined given realization $y^{1,h}$ and s, we simply denote the corresponding values as x^h and y^h, respectively. Then the expectation of $\sigma(s, x, y)$ can be written as

$$\mathbb{E}_{\mathbb{P}}[\sigma(s, x, y)] = \sum\nolimits_{h=1}^{N} \gamma_h \sigma(s, x^h, y^h).$$

Then we have a large integer program M_2.

$$(M_2) \qquad \max \sum_{h=1}^{N} \gamma_h (\sum_{v,i} x^{i,h} - \sum_{v,i} y^{i,h}) \tag{14}$$

$$\text{s.t.} \sum\nolimits_{v \in V} s_v \le k, \tag{15}$$

$$Qs + Tx^h + Wy^h \le b, \; h = 1, \dots, N \tag{16}$$

By solving M_2 , we can get the optimal solution. The number of possible realizations of y^1 is finite but grows exponentially as the size of networks increases. For large graphs, one may resort to *sample average approximation* (SAA), which approximates the expectation using a manageable number of realizations [23]. To obtain a reasonably accurate approximation of our problem, we may end up with a large sample size, which indicates that the SAA problem arising from the above process is also a large-scale programming problem.

To overcome this difficulty, we present a compact relaxation of M_2 from which one can obtain high quality solutions for the d-MCU problem. Consider the constraint of every node $v \in V$ which is directly related with realizations of misinformation early adopters, that is,

$$x_v^{1,h} + y_v^{1,h} \le 1, h = 1, \dots, N.$$

By applying a weighted-averaging technique, we can reduce it into a single constraint $x_v^1 + \sum_{h+1}^{N} \gamma_h y_v^{1,h} \le 1$, which can be further simplified as $x_v^1 + p_v \le 1$. Following the same averaging method, we can have a relaxed mixed integer program M_3.

$$(M_3) \;\; \max \sum_{i=1}^{d} \sum_{v \in V} x_v^i - \sum_{i=1}^{d} \sum_{v \in V} y_v^i$$

$$\text{s.t.} \sum_{v \in V} s_v \le k$$

$$x_v^1 + p_v \le 1, \qquad \forall v \in V$$

$$y_v^1 = p_v, \qquad \forall v \in V$$

$$\text{Constraints (3), (5)-(10) in } \sigma(s, x, y),$$

$$s_v \in \{0, 1\}, x_v^i, y_v^i \in [0, 1]$$

The following theorem shows a bound of M_2 can be obtained from solving M_3.

Theorem 2. *The optimal objective value of M_3 is an upper-bound on the optimal objective value of M_2.*

Proof. To show that the objective of M_3 is an upper-bound on that of M_2, we construct a feasible solution $(\tilde{s}, \tilde{x}, \tilde{y})$ of M_3 that gives an objective equal to the optimal objective of M_2.

Assume $(\hat{s}, \hat{x}^1, \dots, \hat{x}^N, \hat{y}^1, \dots, \hat{y}^N)$ is an optimal solution to M_2. Construct a solution as $(\tilde{s} = \hat{s}, \tilde{x} = \sum_{h=1}^{N} Pr[G^h]\hat{x}^h, \tilde{y} = \sum_{h=1}^{N} \gamma_h \hat{y}^h)$. The objective value of M_2 corresponds to this solution is $\sum\limits_{i=1}^{d} \sum\limits_{v \in V} \sum\limits_{h=1}^{N} \gamma_h x_v^i - \sum\limits_{i=1}^{d} \sum\limits_{v \in V} \sum\limits_{h=1}^{N} \gamma_h y_v^i$ which equals

$$\sum_h \gamma_h \left(\sum_{i=1}^{d} \sum_{v \in V} x_v^{i,h} - \sum_{i=1}^{d} \sum_{v \in V} y_v^{i,h} \right),$$

i.e., the optimal objective of M_2.

Being the first stage variable, $\tilde{s}$ satisfies the budget constraint. We can easily tell that all other constraints are also satisfied since $\tilde{x}$, $\tilde{y}$ are the convex combination of $\hat{x}^h$, $\hat{y}^h$,

$h = 1, 2, \ldots, N$, respectively. Accordingly, we conclude that $(\tilde{s}, \tilde{x}, \tilde{y})$ is a feasible solution to M_3. □

Due to the exponential number of constraints of M_2, it is intractable to obtain optimal solutions. A nice bound from M_3 can provide a helpful tool to evaluate the quality of the proposed algorithm, and this bound helps to determine public expenses for misinformation combat.

VI. Efficient Solution to General Activation

In real social networks, compared with majority-threshold activation, users' acceptance behaviors of information are, however, more complicated. Proposed solution for majority-threshold activation is not computationally efficient in large networks. We want to develop an efficient seed-selection algorithm for general-threshold activation, i.e., acceptance thresholds of information are heterogeneous among users.

Main idea. The main idea of our algorithm is to select the seeds of correct information based on nodes' contribution to maximizing combat score. In d-MCU problem, there exists uncertainty on determining the early adopters of misinformation. To overcome this uncertainty in nodes' contribution evaluation, we simulate the diffusion process of misinformation without correct information and obtain the expected influence weight of misinformation for every node at different time-steps. Later, we iteratively select nodes with the greatest contribution to maximizing combat score into seed set. The contribution of remaining candidates is re-evaluated after each iteration.

Measures. There are two aspects of benefit of adding a new seed. First, by selecting seeding nodes for correct information, we are able to make some nodes A-active. Second, the state changes of some nodes can help block the propagation of misinformation such that many nodes that would otherwise become B-active are saved. We refer this as "deactivation". We define a measure, named *effectiveness*, that incorporates both the promotion on correct information as well as the counteraction of misinformation. Let ε_v denote the effectiveness of node v, and

$$\varepsilon_v = \theta_v^{(a)} + \theta_v^{(b)} + w_v^{(a)} + w_v^{(b)},$$

where all its components are defined as follows.

- $\theta_v^{(a)}$ is the extra weight required to activate node v with correct information;
- $\theta_v^{(b)}$ is the influence wight we need to block in order to deactivate node v if, without any intervention, node v is supposed to become B-active;
- $w_v^{(a)}$ is the influence weight of correct information by adding node v;
- $w_v^{(b)}$ is the influence weight of misinformation that can be blocked by adding node v.

The activation time of nodes is also important. We use r_v^A, r_v^B to denote v's activation round of correct information and misinformation, respectively. For inactive nodes, we set $r_v^A = r_v^B = d+1$ since we are only interested in propagation bounded by d hops. An A-active node (B-active) v is with $r_v^A \leq d$ ($r_v^B \leq d$). The competition between two diffusion processes indicates that only one of r_v^A, r_v^B is in range $[0, d]$. Let f_v^i and g_v^i, respectively, denote the total influence weight on node v from its A-active and B-active neighbors up to i round. In the following part, we give details about how to initialize these measures and update them effectively.

Algorithm 1 Combat Seed Selection (CSS)

```
Data: Graph $G$, budget $k$, probability $\mathbb{P}$, constraint $d$
Result: A seed set $S$ for correct information
Normalization and initialization
for $s$=1 to $k$ do
    Find out node with maximum effectiveness
    $S \leftarrow S \cup \{u^*\}$, initialize a queue $Q \leftarrow \{u^*, r_{u^*}^A, A\}$
    $r_{u^*}^A \leftarrow 0$, $r_{u^*}^B \leftarrow d+1$
    for $v \in N_u^{out}$ do
        $\theta_v^{(a)} \leftarrow \max\{0, \theta_v^{(a)} - w_{uv}^A\}$, $\theta_v^{(b)} \leftarrow \max\{0, \theta_v^{(b)} - w_{uv}^B\}$
    while $Q \neq \emptyset$ do
        $(t, \bar{r}_t^A, X) \leftarrow Q.pop()$;
        if $X = A$ then
            for $w \in N_t^{out}$ do
                for $i = (r_t^A + 1)$ to $\min\{\bar{r}_t^A, d\}$ do
                    $f_w^i \leftarrow f_w^i + w_{tw}^A$;
                for $i = \max\{(r_t^A + 1), r_t^B + 1\}$ to $d$ do
                    $g_w^i \leftarrow g_w^i - w_{tw}^B$;
                if w's A threshold is satisfied no later than B threshold then
                    for $v \in N_w^{out}$ do
                        $\theta_v^{(a)} \leftarrow \max\{0, \theta_v^{(a)} - w_{wv}^A\}$, $\theta_v^{(b)} \leftarrow \max\{0, \theta_v^{(b)} - w_{wv}^B\}$
                    update $r_w^A$
                    if $(w, -, A) \notin Q$ then
                        $Q.push(w, r_w^A, A)$
                else if $\exists i \in [\max\{r_t^A + 1, r_w^B\}, d-1], g_w^i < 1$ then
                    for $v \in N_w^{(out)}$ do
                        $\theta_v^{(b)} \leftarrow \max\{0, \theta_v^{(b)} - w_{wv}^B\}$
                    $r_v^B \leftarrow i + 1$
                    if $(r_w^B < d) \wedge ((w, -, B) \notin Q)$ then
                        $Q.push(w, r_w^B, B)$
        else
            for $v \in N_w^{out}$ do
                for $i = \bar{r}_t^B$ to $\min\{r_t^B - 1, r_w^B - 2\}$ do
                    $g_w^i \leftarrow g_w^i - w_{tw}^B$;
                if $\exists i \in [\max\{r_t^A + 1, r_w^B\}, d-1], g_w^i < \theta_w^B$ then
                    for $v \in N_w^{(out)}$ do
                        $\theta_v^{(b)} \leftarrow \max\{0, \theta_v^{(b)} - w_{wv}^B\}$
                    $r_v^B \leftarrow i + 1$
                    if $(r_w^B < d) \wedge ((w, -, B) \notin Q)$ then
                        $Q.push(w, r_w^B, B)$
```

Initialization. We normalize all edge weights and node thresholds, as discussed in subsection V-A. First, we need to determine the initial values of parameters related with misinformation. We simulate the misinformation diffusion process without any seeds of correct information for L times. In each simulation, we get a realization of misinformation initial adopters, I^B. Given I^B, we can run the diffusion process of misinformation and keep track of how much influence weight every node receives at each time step. Then we have

$g_v^i = \frac{1}{L}\sum_{I^B} g_v^i(I^B)$ for $i = 1, \ldots, d$ and $v \in V$. Let $r_v^B = \text{argmin}_i\, g_v^i \geq \theta_v^B$. If $g_v^d < \theta_v^B$, we set r_v^B to be $d+1$. For each node $v \in V$, we let $\theta_v^{(b)} = \max\{g_v^d - 1, 0\}$. To determine the value of $w_v^{(b)}$, we need to find all connected nodes w with $r_v^B + d(v,w) \leq d$ and $r_v^B \leq d$, denoted as U_v. Then for each node $w \in U_v$, we find all incoming edges from node $u \in U_v \setminus \{w\}$ with $r_u^B > r_w^B$, and let $w_v^{(b)}$ equal the sum of B-weight of these edges. Second, we initialize parameters corresponding to correct information. We let $f_v^i = 0$, $r_v^A = d+1$, and $\theta_v^{(a)} = 1$, for $i = 1, \ldots, d$ and all $v \in V$. When a node v is added into the seed set, we add all v's out-going edges (v,w) into set E_v. If $r_v^B \leq 1$ and $w_{vw}^A \geq 1$, we add all w's outgoing edges into E_v. We keep searching all connected neighbors w with $d(v,w) < d$, and set $w_v^{(b)}$ to be the sum of B-weight of all edges in E_v.

Selection and update. We keep adding the node with the greatest effectiveness among remaining nodes in each round till the given budget is used up. This method favors nodes which help restrain misinformation as well as promote correct information. After that, we need to update measures accordingly, which guarantees that our algorithm makes selection based on the diffusion and the competition between misinformation and correct information within d hops.

If a node v is selected, it causes a chain-reaction, including making inactive connected nodes become active with correct information, accelerating the activation of A-active connected nodes, delaying the activation of misinformation, and even deactivating B-active nodes. To effectively update related measures, we build up a queue Q with elements (v, t_v, I), where I is node state, $I \in \{A, B\}$, and t_v is former activation round, $t_v = r_v^I$. Nodes whose activation state or round are changed will be pushed into Q and used for further updates in a manner similar to the Bellman-Ford shortest-paths algorithm. Every time we pop a node v from Q, if this node has $I = A$, it may trigger the update of its neighbor w's parameters corresponding to both misinformation and correct information, including θ_w^A, θ_w^B, r_w^A, r_w^B, f_v and g_v. If v is inactive or already A-active, we need to check its impact on all out-going neighbors. If neighbor w later changes its state or activation round, and w is not in Q, we push w into Q. If the node poped from Q has $I = B$, we only check its impact on misinformation diffusion. The update process proceeds till Q is empty.

For measures, $w_v^{(a)}$ and $w_v^{(b)}$, we need to recompute their values for each $v \in V$, which is extremely computationally demanding for large OSNs or large time constraint d. Thus, we instead perform a smart update, that is, only update $w_v^{(a)}$ and $w_v^{(b)}$ when it is necessary. Specifically, we construct a max priority queue whose priority is nodes' effectiveness. We pop the top node in this priority queue and recompute its $w_v^{(a)}$ and $w_v^{(b)}$. If its effectiveness is not the highest anymore, we continue to recompute $w_v^{(a)}$ and $w_v^{(b)}$ of the next node. The re-computation and search process stop when the top node still has the highest effectiveness after re-computation. As shown in the experiments, this update procedure greatly reduces running time while producing solutions of high quality.

VII. Experimental Evaluation

A. Experimental Setup

Datasets. To evaluate the effectiveness and efficiency of proposed algorithms, we run experiments on four real datasets: BlogCatalog, Epinions, Livemocha and LiveJournal [24], [25]. Their basic statistics are summarized in Table I.

Dataset	BlogCatalog	Epinions	Livemocha	LiveJournal
Node	10,312	75,879	104,438	2,238,731
Edge	333,983	508,837	2,196,188	14,608,137
Average Degree	32.4	6.7	21	6.5

TABLE I: Statistics of datasets

Setting. Following the widely adopted setting of threshold models [13], [14], [16], [21], edge weights are generated uniformly at random in the range $[0, 1]$. For a node v, its threshold of correct information is selected uniformly at random from $[0, \sum_{u \in N_v^{in}} w_{uv}^A]$. Thresholds of misinformation are set similarly. We are often able to locate a *suspect set* of misinformation adopters which predominantly consists of high out-degree nodes. In our experiments, instead of assuming all nodes are possible misinformation early adopters, we only allow non-zero probability for nodes in suspect set, which is a random subset of top-500 nodes of high out-degree. The size of *suspect set* is 20 in all following experiments. For each node v in suspect set, we assign it a probability p_v selected from uniform distribution in $[0, 1]$ to indicate the probability that v is an initial adopter of misinformation.

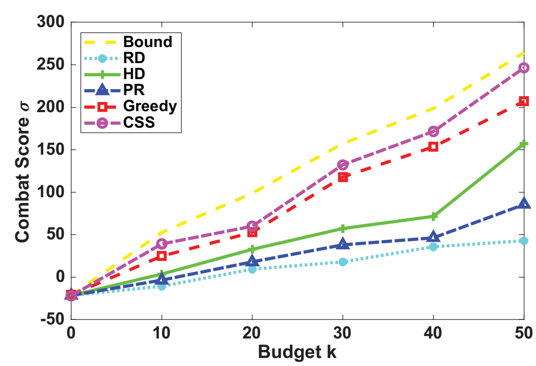
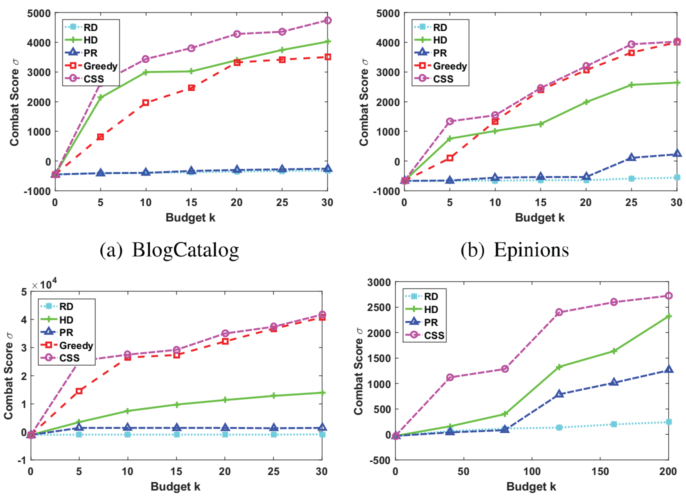


Fig. 2: Combat score under major-threshold activation

Comparison methods. To the best of our knowledge, our work is the first attempt to simultaneously limit misinformation and push out the truth under the uncertainty of misinformation initial adopters. To evaluate the performance of proposed solutions, we compare our algorithms with popular heuristics: Random(RD), High-Degree(HD), PageRank(PR), and Greedy. RD selects seeds randomly. HD favors nodes with high degree (the sum of in-degree and out-degree). PR picks nodes based on their importance. Greedy iteratively selects nodes with largest marginal gain of combat score, which is the state-of-the-art framework in blocking misinformation [15], [26]. Following conventions, 10000 simulations are run for each estimation.

B. Results

Bound. For d-MCU with majority-threshold activation, we can get an upper-bound of combat score by solving the mixed integer program M_3 in subsection V-B. We want to compare results from all algorithms with this bound. We run experiments on a synthetic network with 500 nodes and 5000 edges for $d = 3$. Under majority-threshold activation, only strong exposure can make nodes become active, and thus we set budget $k = 10, 20, 30, 40, 50$. From Fig. 2, one can observe that CSS algorithm can output seeds with combat score that is more closed to the upper-bound than other heuristics.



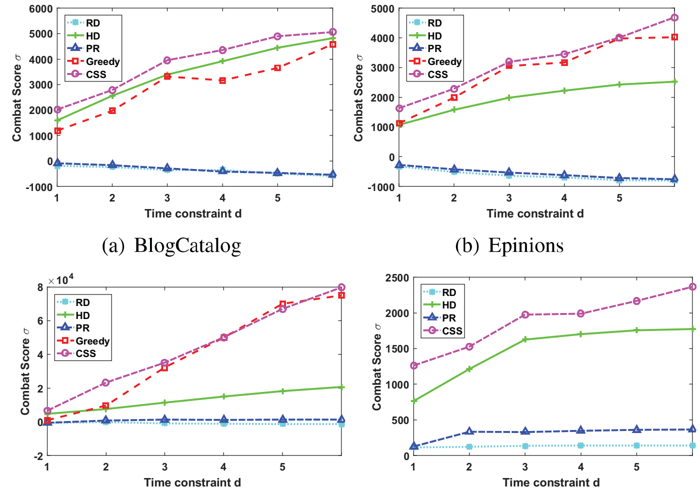
(a) BlogCatalog (b) Epinions

(c) Livemocha (d) LiveJournal

Fig. 3: Combat score as budget increases

Impact of budget. Fig. 3 shows the combat score against different budgets, when time constraint $d = 3$. Greedy appears to be impracticable for large-size networks, and thus is omitted from result comparison in LiveJournal (Fig. 3(d) and Fig. 4(d)). As we observe, the combat score of all algorithms increases as budget increases. HD has better performance than RD and PR. CSS performs consistently well in all networks. For instance, without introducing correct information into BlogCatalog, the expected influence of misinformation is 454, whereas seeding 10 nodes selected by CSS, we can achieve a combat score of 3435. Greedy does not have a good performance for small budget. In Epinions, Greedy's combat score for $k = 5$ is 105, while CSS's is 1344, almost 13 times of Greedy's. Due to activation thresholds, a node probably only activates very few nodes at the very beginning of seed selection. Greedy is unable to consider nodes' contribution for future runs.

Impact of time-constraint. The propagation time is measured by the number of propagation hops d. Since information mostly propagates within a small number of hops [27], we present combat score for d from 1 to 5. One observation is that the combat score of RD and PR does not necessarily increase as time constraint d increases. RD and PR strategy are unlikely promising in misinformation combat. CSS and Greedy have comparative performance for $d = 3, 4, 5$ in Livemocha. But as we show later, CSS requires much less time. For $d = 4$,



(a) BlogCatalog (b) Epinions

(c) Livemocha (d) LiveJournal

Fig. 4: Combat score for different time constraint

CSS can obtain a combat score 50145, which is around half of nodes in Livemocha, while HD's score is 15079. Though LiveJournal has the largest number of nodes, the combat score is relatively small compared with its size. Its sparse connection is one important reason. In our experiments, seeds selected for different time constraints by CSS or Greedy do not have a big overlap, which implies that, in time-sensitive applications, we need to react differently.

Budget	5	10	15	20	25	30
Greedy	698	1506	2295	3090	3912	4622
CSS	16	37	61	87	125	129
Ratio(10^{-2})	2.3	2.5	2.7	2.8	3.2	2.8

TABLE II: Running time for different k in Livemocha

Running time. We want to confirm the efficiency of our algorithm by showing the running time. As the quality of seeds selected by RD, HD, and PR is not comparable with CSS and Greedy, we exclude RD, HD, and PR from the running time comparison. Table II gives the running time measured in minute for Greedy and CSS on Livemocha ($d = 3$), and the ratio of CSS's running time to Greedy's running time. The assessment of node's importance by sampling in Greedy is time-consuming. Our solution can produce comparable solution within much less time, as the the ratio shows.

Time constraint	1	2	3	4	5	6
Greedy	637	689	805	965	969	978
CSS	5	10	16	46	65	58
Ratio(10^{-2})	0.8	1.5	2.0	4.8	6.7	5.9

TABLE III: Running time for different d in Epinions

Table III shows the running time for different time constraints in Epinions. The running time of CSS does not grow linearly as d increases. Note that the required computation of each seed is different in CSS. However, compared with Greedy, it is still very fast. In real applications, we are usually more interested in small time constraints.

VIII. Related Work

Most existing strategies to prevent the spread of virus/misinformation are *node/edge immunization*. Prakash et al. proposed solutions to prevent contagion by minimizing the largest eigenvalue of the graph [28]. Leveraging connection structure, Zhang et al. considered group-level vaccination distribution for both threshold-based and cascade-style contagion [29]. In immunization, the vaccinated nodes can be seen as being removed, but this cannot help disseminate correct information. A growing amount of efforts have been devoted to *Competitive propagation*. Kostka et al. studied the influence maximization problem for competing rumors utilizing game theory [30]. Borodin proposed several extensions of linear-threshold model to represent competition between two entities [17]. He et al. [21] proposed solutions to block misinformation based on a local directed acyclic graph structure. Given limited budget, Carnes et al. tried to maximize the follower's influence in the presence of adopters of a competing product [31]. However: they either consider maximizing one campaign in the presence of a competing campaign or focus on reducing the influence of such a campaign. We instead aim to combat misinformation by jointly considering both aspects.

IX. Conclusion

In this paper, we introduce the misinformation combat problem d-MCU. We consider the uncertainty in identifying misinformation initial adopters and prove the #P-hardness of the computation of combat score. For majority-threshold activation, we provide a solution based on stochastic programming. In addition, we propose efficient strategy to disseminate correct information, which greatly saves running time and retains solution quality. Extensive experiments validate the efficiency of our algorithms. We can expect at least two new directions. First, the combination of internal and external influence is a promising strategy for the fight against misinformation. Second, heterogeneous time delays of information propagation are considered.

Acknowledgement

This work is partially supported by the NSF CNS-1443905 and CCF-1422116.

References

[1] M. Del Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi, "The spreading of misinformation online," *PNAS*, vol. 113, no. 3, pp. 554–559, 2016.

[2] L. Zeng, K. Starbird, and E. S. Spiro, "Rumors at the speed of light? modeling the rate of rumor transmission during crisis," in *HICSS*, 2016.

[3] F. Jin, W. Wang, L. Zhao, E. R. Dougherty, Y. Cao, C.-T. Lu, and N. Ramakrishnan, "Misinformation propagation in the age of twitter." *IEEE Computer*, vol. 47, no. 12, pp. 90–94, 2014.

[4] S. Lewandowsky, U. K. Ecker, C. M. Seifert, N. Schwarz, and J. Cook, "Misinformation and its correction continued influence and successful debiasing," *Psychological Science in the Public Interest*, vol. 13, no. 3, pp. 106–131, 2012.

[5] L. Howell *et al.*, "Digital wildfires in a hyperconnected world," *WEF Report*, vol. 3, pp. 15–94, 2013.

[6] K. K. Kumar and G. Geethakumari, "Detecting misinformation in online social networks using cognitive psychology," *Human-centric Computing and Information Sciences*, vol. 4, no. 1, p. 1, 2014.

[7] S. Kwon, M. Cha, K. Jung, W. Chen, and Y. Wang, "Prominent features of rumor propagation in online social media," in *ICDM*. IEEE, 2013.

[8] E. Seo, P. Mohapatra, and T. Abdelzaher, "Identifying rumors and their sources in social networks," in *SPIE defense, security, and sensing*. International Society for Optics and Photonics, 2012, pp. 83 891I–83 891I.

[9] P. Basaras, D. Katsaros, and L. Tassiulas, "Dynamically blocking contagions in complex networks by cutting vital connections," in *ICC*. IEEE, 2015.

[10] L. Bode and E. K. Vraga, "In related news, that was wrong: The correction of misinformation through related stories functionality in social media," *Journal of Communication*, vol. 65, no. 4, pp. 619–638, 2015.

[11] H. Zhang, A. Kuhnle, H. Zhang, and M. T. Thai, "Detecting misinformation in online social networks before it is too late," in *ASONAM*. IEEE, 2016, pp. 541–548.

[12] H. Zhang, M. A. Alim, M. T. Thai, and H. T. Nguyen, "Monitor placement to timely detect misinformation in online social networks," in *ICC*. IEEE, 2015, pp. 1152–1157.

[13] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *SIGKDD*, 2003.

[14] W. Chen, Y. Yuan, and L. Zhang, "Scalable influence maximization in social networks under the linear threshold model," in *ICDM*, 2010.

[15] C. Budak, D. Agrawal, and A. El Abbadi, "Limiting the spread of misinformation in social networks," in *WWW*, 2011.

[16] H. Zhang, M. A. Alim, X. Li, M. T. Thai, and H. T. Nguyen, "Misinformation in online social networks: Detect them all with a limited budget," *ACM Transactions on Information Systems*, vol. 34, no. 3, p. 18, 2016.

[17] A. Borodin, Y. Filmus, and J. Oren, "Threshold models for competitive influence in social networks," in *International Workshop on Internet and Network Economics*. Springer, 2010, pp. 539–550.

[18] N. Chen, "On the approximability of influence in social networks," *SIAM Journal on Discrete Mathematics*, vol. 23, no. 3, pp. 1400–1415, 2009.

[19] Z. Lu, W. Zhang, W. Wu, J. Kim, and B. Fu, "The complexity of influence maximization problem in the deterministic linear threshold model," *Journal of combinatorial optimization*, vol. 24, no. 3, pp. 374–378, 2012.

[20] T. Soma, N. Kakimura, K. Inaba, and K.-i. Kawarabayashi, "Optimal budget allocation: Theoretical guarantee and efficient algorithm," in *ICML*, 2014.

[21] X. He, G. Song, W. Chen, and Q. Jiang, "Influence blocking maximization in social networks under the competitive linear threshold model." in *SDM*, 2012.

[22] D. Peleg, "Local majority voting, small coalitions and controlling monopolies in graphs: A review," in *SIROCCO*, 1997, pp. 152–169.

[23] A. J. Kleywegt, A. Shapiro, and T. Homem-de Mello, "The sample average approximation method for stochastic discrete optimization," *SIAM Journal on Optimization*, vol. 12, no. 2, pp. 479–502, 2002.

[24] R. Zafarani and H. Liu, "Social computing data repository at ASU," 2009. [Online]. Available: http://socialcomputing.asu.edu

[25] "Stanford large network dataset collection." [Online]. Available: http://snap.stanford.edu/data/index.html

[26] G. A. Tong, W. Wu, L. Guo, D. Li, C. Liu, B. Liu, and D.-Z. Du, "An efficient randomized algorithm for rumor blocking in online social networks," in *INFOCOM*, 2017.

[27] M. Cha, A. Mislove, and K. P. Gummadi, "A measurement-driven analysis of information propagation in the flickr social network," in *WWW*, 2009.

[28] B. A. Prakash, L. Adamic, T. Iwashyna, H. Tong, and C. Faloutsos, "Fractional immunization in networks," in *SDM*, 2013.

[29] Y. Zhang, A. Adiga, A. Vullikanti, and B. A. Prakash, "Controlling propagation at group scale on networks," in *ICDM*, 2015.

[30] J. Kostka, Y. A. Oswald, and R. Wattenhofer, "Word of mouth: Rumor dissemination in social networks," in *SIROCCO*. Springer, 2008, pp. 185–196.

[31] T. Carnes, C. Nagarajan, S. M. Wild, and A. Van Zuylen, "Maximizing influence in a competitive social network: a follower's perspective," in *ICEC*, 2007.