Uncertainty-Aware Influence Maximization: Enhancing Propagation in Competitive Social Networks with Subjective Logic

Qi Zhang*, Lance M.Kaplan†, Audun Jøsang‡, Dong Hyun Jeong§, Feng Chen¶, Jin-HeeCho*
* Department of Computer Science, Virginia Tech, USA
†US DEVCOM Army Research Laboratory, USA
‡Department of Informatics, University of Oslo, Norway
§Department of Computer Science, University of the District of Columbia, USA
¶Department of Computer Science, University of Texas at Dallas, USA

2024 IEEE International Conference on Big Data (BigData) | 979-8-3503-6248-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/BigData62323.2024.10825012

***Abstract*—The Competitive Influence Maximization (CIM) problem involves entities competing to maximize influence in online social networks (OSNs). While Deep Reinforcement Learning (DRL) methods have shown promise, most assume binary user opinions and overlook behavioral factors. We introduce DRIM, a novel DRL-based CIM framework using Subjective Logic (SL) to incorporate user preferences and uncertainty, optimizing seed selection to spread true information while countering false information. DRIM's Uncertainty-based Opinion Model (UOM) provides a realistic representation of user opinions. Results demonstrate that UOM maintains over 80% true influence against advanced misinformation, and DRIM outperforms state-of-the-art methods by up to 45% in influence and 77% in speed. DRIM also excels in limited-resource scenarios, networks with 10% invisibility, and when users are inclined to doubt true information.**

***Index Terms*—Competitive influence maximization, deep reinforcement learning, uncertainty, opinion models**

I. Introduction

Online Social Networks (OSNs) serve as a platform for information sharing and opinion shaping, where competitive environments arise in contexts like service selection and voting. Known as the Competitive Influence Maximization (CIM) problem, this involves entities, such as political parties or companies, strategically selecting seed nodes to maximize influence. Our research addresses a scenario with two competing forces: the *true party*, promoting accurate information, and the *false party*, spreading misinformation that can cause reputation damage, financial loss, and biased public opinion.

To counter misinformation effectively, we explore strategies that empower the *true party* to limit false content spread, enhancing the visibility of truthful information in OSNs. While Deep Reinforcement Learning (DRL) has shown promise in CIM for real-time strategy optimization, current approaches often simplify user opinions as static and binary, missing nuances in user behavior and uncertainty. Our approach fills this gap by modeling user opinion dynamics to reflect uncertainty and conflicting decision-making, leading to a more realistic representation of influence in OSNs.

We propose a DRL-based CIM framework, *DRIM*, to optimize the spread of true information and counteract false information by identifying influential seed users, adapting to the dynamic and uncertain nature of user opinions. Our **key contributions** are: (1) We integrate *Subjective Logic* (SL) [8] into CIM to capture the uncertain and evolving nature of user opinions, moving beyond binary opinion models commonly seen in prior CIM studies [1, 2, 5, 15], and enhancing our understanding of opinion shifts in OSNs. (2) We develop a dual-agent DRL model to simulate true and false information spread, advancing strategic depth beyond current single-agent models [5, 14, 15]. (3) We assess the Uncertainty-aware Opinion Model (UOM) for its ability to capture early false information dominance, which complicates correction efforts, and demonstrate UOM's effectiveness in boosting engagement with accurate content and reducing misinformation spread. (4) Finally, we analyze partial network observability impacts on CIM, especially regarding non-binary opinions with UOM, addressing critical gaps in understanding visibility effects on CIM outcomes.

We use the term *false information* rather than *misinformation* (unintentionally shared) or *disinformation* (intentionally misleading) for clarity, as distinctions between these terms are beyond our research scope.

II. Related Work

The Influence Maximization (IM) problem [6] seeks to maximize influence in OSNs by selecting an optimal set of seed nodes to reach the largest number of nodes through a specified propagation process. Kempe et al. [11] formalized IM as a discrete stochastic optimization problem, solidifying its theoretical basis. CIM extends IM by involving multiple entities competing to maximize influence within a single OSN, with various methodologies proving effective in tackling this challenge.

979-8-3503-6248-0/24/$31.00 ©2024 IEEE

A. Propagation Models

Propagation models like Independent Cascade (IC) [12] and Linear Threshold (LT) [7] are widely used to analyze network influence [3, 16, 17, 19]. The IC model activates nodes to independently influence neighbors, while the LT model activates nodes based on cumulative neighbor influence exceeding a threshold. Although effective, both models use static probabilities and thresholds, limiting real-world applicability. Our study improves these models by incorporating dynamic opinion dynamics, assessing their robustness against false information spread.

B. DRL-based CIM

The STORM framework introduced a STrategy-Oriented ReinforceMent-Learning-based approach, pioneering RL in multi-round CIM within the competitive LT model [15] by enabling dynamic strategy selection based on evolving opinions and competitor actions. This was enhanced with a DRL-based CIM framework using Deep-Q Learning (DQN) and community detection for multi-round seed selection [5]. Subsequent work adapted STORM's reward and action spaces for unknown topologies by adding network exploration actions [1].

Previous work [5, 15] often assumed full network knowledge, unrealistic in practice, and used binary, static opinion models [1, 2]. Our approach addresses these limitations by incorporating detailed opinion dynamics and uncertainties in user opinions and network structure, resulting in a more realistic CIM framework.

III. Problem Statement

An OSN is modeled as an unweighted, undirected graph $G = (V, E)$, with V as users and E as relationships. There are *true party* (TP) and *false party* (FP) aiming to maximize their influence within G. Each of them has a central decision maker (CDM) to select seed users $\mathbf{S}$ ($\mathbf{S}^{TP}$ or $\mathbf{S}^{FP}$) to spread their respective information types. At each moment, given the current users' opinion status in OSN, DRL is employed to select action $\boldsymbol{A}$ to pick seed nodes for maximizing influence. Each party's node selection process will follow the procedures and opinion models described in Section IV. We define each party's influence by the number of users who believe either true or false information. Since users' opinions are not binary, we will explain how to define whether a user believes true or false information after introducing the SL-based opinion model in Section IV-A.

IV. Proposed CIM Framework

A. Subjective Logic (SL)-based Opinion Formulation

In binary logic, a user's opinion toward something must be either 100% believed in or not, which barely happens. To better reflect real-world opinions, we incorporate uncertainty to model users' degrees of uncertainty about beliefs. To represent opinions on true or false information, we use SL's binomial model to capture users' opinions in OSNs. An opinion ω is defined as $\omega = (b, d, u, a)$, where belief b reflects agreement with true information, disbelief d captures agreement with false information (or disagreement with true information), and uncertainty u accounts for a lack of sufficient evidence. Each component is a real number within the range $[0, 1]$, with the condition that $b + d + u = 1$. These components are defined as:

$$b = \frac{r}{r+s+W}\ , d = \frac{s}{r+s+W}\ , u = \frac{W}{r+s+W}\ , \quad (1)$$

where r and s are the numbers of evidence to support b and d, respectively, and W refers to the number of uncertain evidence that cannot be judged as true or false, supporting neither b nor d.

The base rate a, a real number in $[0, 1]$, represents the prior belief favoring true information, with $1 - a$ indicating disbelief (i.e., favoring false information). While this paper focuses on two opinions, the approach can be extended to multiple parties.

To make decisions (e.g. which product to buy), users incorporate uncertainty into their decisions by using the *projected probability* in SL, such as the projected belief or disbelief, denoted by $P(b)$ and $P(d)$:

$$P(b) = b + a \times u,\ P(d) = d + (1 - a) \times u, \quad (2)$$

where $P(b) + P(d) = 1$ and $a \times u + (1 - a) \times u = u$. Since users choose between b and d, they interpret uncertainty based on their prior belief, $\boldsymbol{a} = \{a, 1 - a\}$, where $a + (1 - a) = 1$. To be specific, user i is considered part of the *true party* (TP) if $P(b_i) > 0.5$ and is deemed to believe false information if $P(d_i) > 0.5$. When a user i has $P(b_i) = P(d_i)$ under uniform prior belief (base rate), the user is not counted toward TP or FP. However, such cases are extremely rare, so the total number of users in TP or FP is nearly equal to the total number of users in a given network.

We consider two types of uncertainty in the user's opinion [9]: *vacuity uncertainty* and *dissonance uncertainty*. *Vacuity* refers to uncertainty caused by a lack of evidence, while *dissonance* indicates uncertainty due to conflicting evidence. The *vacuity uncertainty* mass is measured by u, and the *dissonance uncertainty* mass is defined as:

$$\boldsymbol{b}^{\text{Diss}} = (b + d) \cdot \text{Bal}(b, d), \quad (3)$$

$\text{Bal}(b, d)$ is the relative mass balance between b and d:

$$\text{Bal}(b, d) = 1 - \frac{|b - d|}{b + d}. \quad (4)$$

We incorporated uncertainty estimates, such as *vacuity* and *dissonance*, to develop a UOM in Section IV-C.

B. User Types

We categorize users in OSNs into three types:

- *True information propagators* (TIPs) are the seed nodes chosen by *true party* (TP), their opinion is initialized as $\omega = (b, d, u, a) = (b \to 1, d \to 0, u \to 0, a = 1)$. This opinion implies TIPs have a strong belief in true information (b is close to 1), while they lack belief in false information (d is close to 0).
- *False information propagators* (FIPs) are the seed nodes chosen by *false party* (FP), their opinion is initialized as $\omega = (b, d, u, a) = (b \to 0, d \to 1, u \to 0, a = 0)$. This means FIPs have a strong belief in false information (d is close to 1) while having a lack of belief in true information (b is close to 0).
- *Legitimate users* in OSNs are regular users with highly uncertain opinions, initialized as $\omega = (b, d, u, a) = (b \to 0, d \to 0, u \to 1, a = 0.5)$.

Since TIPs and FIPs hold strong beliefs, they do not change their opinions, while *legitimate users* update their opinions by adjusting (b, d, u, a) based on Eq. (5).

C. Opinion Models

Each user i's opinion $\omega_i = (b_i, d_i, u_i, \boldsymbol{a}_i)$ will be affected by their behaviors, which consists of three main components: opinion updating, sharing, and reading.

1) **Opinion Updating:** When user i encounters user j and reads user j's information (i.e., ω_j), user i will update the opinion. We leverage the *consensus operator* in SL [8] to calculate i's new opinion. The consensus operator determines how two users can reach an agreement based on their current opinions. User i's opinion after interacting with user j is updated by combining its own belief weighted by j's uncertainty and j's belief weighted by its uncertainty. Similarly, i's disbelief is updated using its disbelief weighted by j's uncertainty and j's disbelief weighted by its uncertainty. i's uncertainty after the interaction is reduced to the product of both users' uncertainty levels. Opinion updates occur as long as neither party's uncertainty reaches zero, which would indicate complete confidence. Following this process, user i's updated opinion, ω_i', is given by:

$$\begin{aligned} \omega_i' &= \omega_i \oplus \omega_{i\otimes j} = (b_i', d_i', u_i', a_i) \\ &= (b_i \oplus b_{i\otimes j}, d_i \oplus d_{i\otimes j}, u_i \oplus u_{i\otimes j}, a_i \oplus a_{i\otimes j}), \end{aligned} \tag{5}$$

where ω_i is user i's previous opinion, and $\omega_{i\otimes j}$ represents j's discounted opinion according to the degree that i accepts j's opinion. Here we introduce a *trust filter*, c_i^j, to discount j's opinion because user i accepts user j's opinion to the extent that i trusts j. Each component of $\omega_{i\otimes j} = (b_{i\otimes j}, d_{i\otimes j}, u_{i\otimes j}, a_{i\otimes j})$ is estimated by:

$$\begin{aligned} b_{i\otimes j} &= c_i^j b_j, \ d_{i\otimes j} = c_i^j d_j, \\ u_{i\otimes j} &= 1 - c_i^j(1-u_j), \ a_{i\otimes j} = a_j. \end{aligned} \tag{6}$$

Finally, user i's new opinion is calculated by applying the consensus operator to its original opinion with discounted j's opinion. The $\omega_i \oplus \omega_{i\otimes j}$ is obtained by:

$$\begin{aligned} b_i \oplus b_{i\otimes j} &= \frac{b_i(1-c_i^j(1-u_j)) + c_i^j b_j u_i}{\beta}, \\ d_i \oplus d_{i\otimes j} &= \frac{d_i(1-c_i^j(1-u_j)) + c_i^j d_j u_i}{\beta}, \\ u_i \oplus u_{i\otimes j} &= \frac{u_i(1-c_i^j(1-u_j))}{\beta}, \\ a_i \oplus a_{i\otimes j} &= \frac{(a_i-(a_i+a_j)u_i)(1-c_i^j(1-u_j)) + a_j u_i}{\beta - u_i(1-c_i^j(1-u_j))}, \\ \beta &= 1 - c_i^j(1-u_i)(1-u_j) \neq 0. \end{aligned} \tag{7}$$

For how to calculate the *trust filter* c_i^j, we introduce the **uncertainty-based trust opinion model (UOM)**. This model represents users who seek new information, particularly when they lack sufficient evidence to form a conclusion [4]. UOM employs uncertainty-based trust, and the uncertainty *trust filter* c_i^j is defined as:

$$c_i^j = (1-u_i)(1-u_j). \tag{8}$$

To ensure non-zero β (non-zero denominator) in Eq. (7), we use the *vacuity maximization* technique [8] to increase u_i when a user's uncertainty is very small (i.e., near 0) but they still cannot decide due to high conflicting evidence. Given two thresholds T_v and T_d, when $u_i < T_v$ (low vacuity with sufficient evidence) and dissonance $\boldsymbol{b}_i^{\text{Diss}} > T_d$ (high uncertainty due to conflicting evidence), we update user i's opinion to a *vacuity-maximized opinion*, denoted as $\ddot{\omega}_i = (\ddot{b}_i, \ddot{d}_i, \ddot{u}_i, \boldsymbol{a}_i)$, with $\ddot{b}_i$, $\ddot{d}_i$, and $\ddot{u}_i$ calculated as:

$$\begin{aligned} \ddot{u}_i &= \min\Big[\frac{P(b_i)}{a_i}, \frac{P(d_i)}{1-a_i}\Big], \\ \ddot{b}_i &= P(b_i) - a_i \cdot \ddot{u}_i, \ \ \ddot{d}_i = P(d_i) - (1-a_i)\cdot \ddot{u}_i, \end{aligned} \tag{9}$$

where $\ddot{u}_i$ allows user i to continue to accept new information and update its opinion. We set $T_v = 0.01$ and $T_d = 0.6$ as thresholds to effectively maximize each party's influence based on our experimental analysis.

2) **Opinion Reading:** Following [4], we define a user's reading probability, P_r, as the frequency they read information from neighbors. For each user, P_r is uniformly randomly assigned from 1 (multiple times per day), 0.5 (daily), 0.25 (weekly/monthly), or 0.1 (rarely) to model different reading habits.

3) **Opinion Sharing:** Similarily, each user shares an opinion ω with friends based on a sharing probability P_s. Following the distribution from [4], P_s is uniformly randomly assigned from 1 (always/mostly), 0.5 (half the time), 0.25 (sometimes), and 0.1 (rarely).

Since reading motivates opinion updates [10], we assume users share only after reading. User i shares its current opinion ω_i, not the original 'True' or 'False' information from the *true party* (TP) or *false party* (FP). Only seed nodes, *True* and *False information propagators* (TIPs and FIPs), share the original information.

D. Partially Observable Network

Following [18], we define the partially observable network $G' = (V, E')$ of an undirected graph $G = (V, E)$, where $E' \subset E$. This means only a portion of the edges are visible to the DRL agent when observing the state in OSNs. E' is randomly chosen from E.

E. DRL-based Seed Set Selection Process

We use DRL to optimize strategy selection for maximizing influence spread in a multi-round process. In each round, the FP selects a seed node (turns it into an FIP), and then the FIP starts sharing its opinion with its neighbors and propagates further until all nodes are reached. The TP next selects an TIP to propagate true information. Information spreads through the network via *Breadth-First Search*, with nodes deciding to update and propagate based on their reading and sharing probability. An episode consists of T rounds, matching the number of seed nodes.

1) **States:** At each round t, the state s_t is defined as:

$$s_t = \left\{ \sum_{i,j \in \mathcal{U}} e_{i,j}, \max_{i \in \mathcal{U}} \deg_i \right\}, \tag{10}$$

where $\sum_{i,j \in \mathcal{U}} e_{i,j}$ counts the edges among *free nodes*, and $\max_{i \in \mathcal{U}} \deg_i$ gives the highest degree among them. *Free nodes* are defined as $\{j \in \mathcal{U} | u_j \geq 0.5\}$, meaning user j has high uncertainty and has not aligned with either party. Here, u_j refers to the user's 'uncertainty' or 'vacuity' (see Eq. (1)).

2) **Actions:** The action space at round t is $\mathbf{a}_t = \{a_t^{AF}, a_t^{BF}, a_t^{SGF}, a_t^{CF}\}$, focusing on user behavior and centrality in the OSN, where the action space includes:

- *Active First (AF, a_t^{AF})* prioritizes the most active user, determined by the highest $P_r \times P_s$, where P_r and P_s represent the user's reading and sharing probabilities.
- *Blocking First (BF, a_t^{BF})* targets neighbors of the opponent's party with the highest *free degree* (connected to *free nodes*. A user belongs to the TP if $P(b_i) > 0.5$, or the FP if $P(d_i) > 0.5$ (see Eq. (2)).
- *SubGreedy First (SGF, a_t^{SGF})* selects the node with the most neighbors within d-hops [5, 15]. We set $d = 2$ to balance efficiency and effectiveness.
- *Centrality First (CF, a_t^{CF})* selects the user with the highest degree centrality.

3) **Rewards:** We use instant rewards to enhance learning process. At each round t, the rewards for each party are the net change in users aligned with each party:

$$R_t^{TP} = n_t^{TP} - n_{t-1}^{TP}, \quad R_t^{FP} = n_t^{FP} - n_{t-1}^{FP}. \tag{11}$$

TABLE I: KEY PARAMETERS AND DEFAULT VALUES

Param.	Meaning	Def. Val.
T	Number of information propagation rounds	50
T_v	Vacuity threshold in UOM	0.01
T_d	Dissonance threshold in UOM	0.6
lr_a	Learning rate in the actor-network	0.0003
lr_c	Learning rate in the critic network	0.001
K_{epo}	The number of epochs used in one PPO update	80
ϵ	Clipping parameter value used in PPO	0.2
γ	Discount factor in DRL's reward function	0.95
p^{TP}	Number of true information propagation by the true party's seed nodes	2
p^{FP}	Number of false information propagation by the false party's seed nodes	1
d	Value used in the Subgreedy strategy	2
a	Prior belief of legitimate user	0.5

The accumulated reward over an episode is:

$$R_T^{TP} = \sum_{t=1}^{T} \gamma^{T-t+1} R_t^{TP}, \quad R_T^{FP} = \sum_{t=1}^{T} \gamma^{T-t+1} R_t^{FP}, \tag{12}$$

where γ is the discount factor.

Fig. 1 illustrates the proposed DRL-based CIM (DRIM) framework. Given an OSN, two parties compete to gain more users believing in their opinion. Initially, all users are legitimate and neutral. In each round, given the current state in OSN, FP's central decision maker (CDM) selects a seed node according to the action distribution output from the policy network (i.e., criteria to maximize information influence). The selected node (FIP) propagates opinions based on the opinion model in Section IV-C (UOM). Then, TP's CDM chooses an action to select a seed node for maximizing its information influence. Once terminated (picked up to predefined seeds or no more seeds can be selected), TP's influence is evaluated by counting how many nodes believe in TP.

V. EXPERIMENT SETUP

Our DRL framework employs Proximal Policy Optimization (PPO) [20] for optimal seed node selection, leveraging PPO's efficiency in reusing data to reduce computational complexity, critical for CIM's long data collection time. TP uses a trained DRL model for node selection, while FP chooses among six strategies: DRL, Active First (AF), Blocking First (BF), SubGreedy First (SGF), Centrality First (CF), and Random. We train six DRL agents for TP, each paired with one FP strategy, and test these configurations in Section VI.

We initialize SL-based opinions for legitimate users using the mapping rule in Eq. (1), with $(r, s, W) = (1, 1, 101)$, indicating high uncertainty. TIPs start with $(r, s, W) = (100, 1, 2)$, reflecting strong confidence in true information, while FIPs are initialized at $(r, s, W) = (1, 100, 2)$, showing confidence in false information.

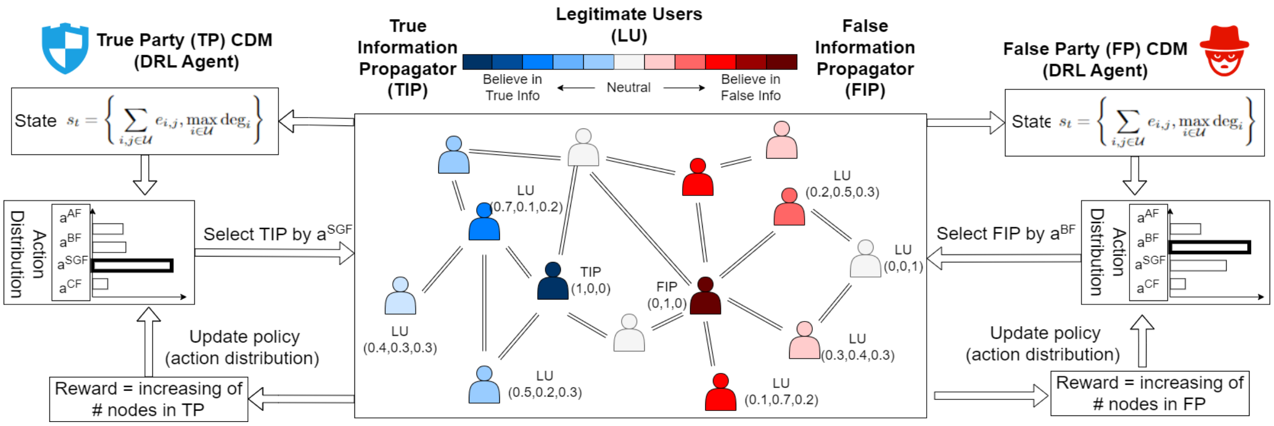


Fig. 1: **Overview of DRIM**: The center shows an SL-based competitive network with color gradients for belief strength – lighter colors indicate higher uncertainty, white for neutral users. Blue shades represent true information supporters (darkest as TIP), red shades indicate false information supporters (darkest as FIP). On each side, DRL-based CDMs receive the OSN state, use a policy network to select the highest-probability action for node selection, and update based on reward feedback.

FP and TP alternately pick 50 seeds and propagate (FP goes first, followed by TP). This models real false information mitigation scenarios, where true information counters ongoing false information. TP propagates twice after each seed selection, while FP propagates once. Since human judgment and public awareness are crucial in combating false information, true information is assumed to achieve more effective propagation in OSNs. All experiment results are averaged over 50 runs.

All experiments are conducted on an HPE Apollo 6500 system with AMD EPYC 7742 chips, running at a base frequency of 2.25 GHz and a boost up to 3.4 GHz. Table I summarizes the key design parameters, their meanings, and default values. For all performance analyses, except those examining varying network observability, we assume perfect network observability.

Datasets: We use the Facebook social circles dataset [13] to validate our framework. It has 4039 nodes and 88234 edges. This dataset is well-suited for studying CIM, as it represents real users and interactions, and models how individuals share ideas, opinions, and information through conversations, posts, or recommendations, capturing social influence dynamics. We tested on extra datasets of various sizes and densities to verify the generalizability of our framework, and they both showed similar results as the Facebook network. Due to the page limitation, we are not sharing in this paper.

Metrics: The performance of each framework is measured by the number of users aligned with the TP (see Eq. (2)). Users with $P_i(b_i) > 0.5$ are classified under TP, indicating TP's influence, denoted as n^{TP}. Algorithmic efficiency is evaluated by the simulation's running time per round in Table II.

Comparing Schemes: To demonstrate the effectiveness of our framework, we compare its performance and running time against existing DRL-based CIM frameworks. Unlike [3, 5, 15, 19], our approach uses an SL-based dynamic opinion model, allowing opinions to evolve through interactions. For a fair comparison with existing state-of-the-art approaches, all models operate under the same conditions, including identical opinion models (UOM), same seed node selection setting (FP moves first, equal information propagation times, and $T = 50$ seed nodes), and identical opinion update mechanisms as described in Section IV. While adapting these schemes may affect their original performance, this standardization is essential for evaluating their performance in the realistic scenarios addressed in our study. We evaluate the performance of the following CIM algorithms: (1) **DRIM-A**: Our proposed framework DRIM. 'A' refers to the AF strategy (see Section IV-E), we append '-A' in contrast to DRIM-NA. (2) **DRIM-NA**: Excludes the AF strategy (No-AF) from the action space for node selection, as we want to see the effect of AF in DRIM. (3) **STORM** [15]: Originally based on binary opinions for node occupation, we adapt it by defining *free nodes* ($\{j | u_j \geq 0.5\}$) as unoccupied. We also merge max-weight and max-degree actions since our datasets are unweighted graphs. (4) **C-STORM** [5]: Enhances STORM by introducing a preliminary community detection step for optimal seed selection. We adapt it using our opinion model and free nodes definition similar to STORM.

VI. Numerical Analysis & Results

Effect of DRL-based TP's Influence Under Various Strategies Taken by FP: Fig. 2 compares four DRL-based TP agents against FP across six seed selection strategies, with the x-axis showing FP's strategy and the y-axis the percentage of nodes aligned with TP (normalized). In Fig. 2(a), DRIM-A and DRIM-NA out-

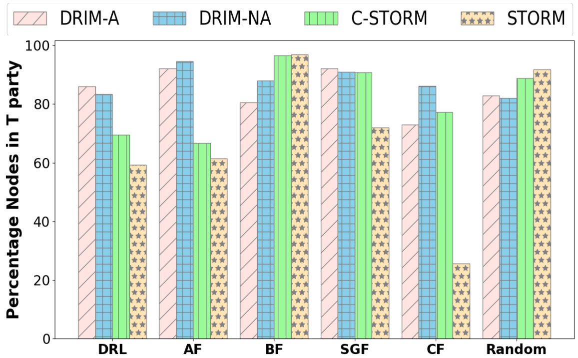


Fig. 2: TP's influence under various CIM algorithms

perform C-STORM and STORM when FP uses DRL, AF, SGF, or CF. When the FP uses the DRL-based strategy, in the Facebook network, DRIM-A achieves 86.01% TP influence, DRIM-NA 83.33%, while C-STORM and STORM reach only 69.42% and 59.21%, showing up to 45.26% higher influence. Lower performance against BF and Random suggests our schemes are optimized for proactive opponents.

We may perceive that DRIM-A and DRIM-NA perform "worse" when FP uses BF or Random. However, they maintain over 80% influence, showing stable performance. While C-STORM and STORM may surpass our schemes under BF and Random, they are ineffective for spreading false information. Moreover, DRIM-A and DRIM-NA outperform them when FP employs strategic methods (e.g., DRL, AF, CF, SGF), highlighting DRIM's effectiveness to combat false information in OSNs.

Sensitivity Analyses: In this experiment, TP and FP both use DRL-based agents to select seed nodes, allowing us to assess TP's performance against "smart" opponents.

1) **Effect of Varying the Number of Information Propagations by TIPs:** Fig. 3(a) explores the impact of increasing TP's information propagations (IPs) (from 1 to 5) while FP remains at one propagation per seed. As expected, all schemes show greater influence with more IPs. Notably, DRIM-A and DRIM-NA perform best with 1-2 IPs per round, effectively countering false information with minimal resources. When IPs exceed 2, the increase plateaus, supporting our choice of 2 as the default for optimizing resources and maximizing influence.

2) **Effect of Varying the Degree of Network Observability:** As shown in Fig. 3(b), all schemes naturally gain more influence with increased network visibility. Notably, DRIM-A and DRIM-NA begin outperforming once visibility reaches 85%, demonstrating their ability to adapt to uncertain environments.

3) **Effect of Varying Users' Prior Belief:** Prior belief reflects the initial likelihood of accepting true information. A higher prior belief increases TP's influence, as shown in Fig. 3(c). Notably, even with low ($a < 0.5$) or neutral ($a = 0.5$) prior beliefs, DRIM-A and DRIM-NA outperform C-STORM and STORM, demonstrating effectiveness even when initial user beliefs do not favor true information.

TABLE II: SIMULATION RUNNING TIME (IN SEC.) OF THE CONSIDERED CIM ALGORITHMS

Algo.	DRIM-A	DRIM-NA	C-STORM	STORM
FB	5.456	5.095	10.445	23.812

In summary, DRIM-A and DRIM-NA are resource-efficient and more effective in partially observable networks. They successfully combat false information by maximizing the spread of true information, even when users initially lean toward false information.

Running Time Analysis of CIM Algorithms: Table II shows DRIM-A and DRIM-NA outperform C-STORM and STORM in running time. We measure the duration of selecting a node and completing information propagation from the seed to all other nodes. These averages were obtained from 2,500 node selection and propagation cycles on the the dataset. DRIM is 47.8% faster than C-STORM and 77.1% faster than STORM in the Facebook network. Thus, our DRIM-based approach is notably faster, especially in denser networks.

VII. CONCLUSIONS

This work introduces a deep reinforcement learning framework enhanced with Subjective Logic to improve competitive influence maximization (CIM) in OSNs by accounting for uncertain opinions and user preferences. Unlike binary-opinion CIM models, our Uncertainty-based Opinion Model (UOM) provides a nuanced representation of user opinions. The proposed frameworks, DRIM-A and its variation DRIM-NA, demonstrate superior efficiency and effectiveness, outperforming C-STORM and STORM by up to 23.9% and 45.26% in influence spread while using 47.8% and 77.1% less time.

Experimental results indicate UOM enhances true information spread, even among low-engagement users, highlighting the importance of targeting active users.

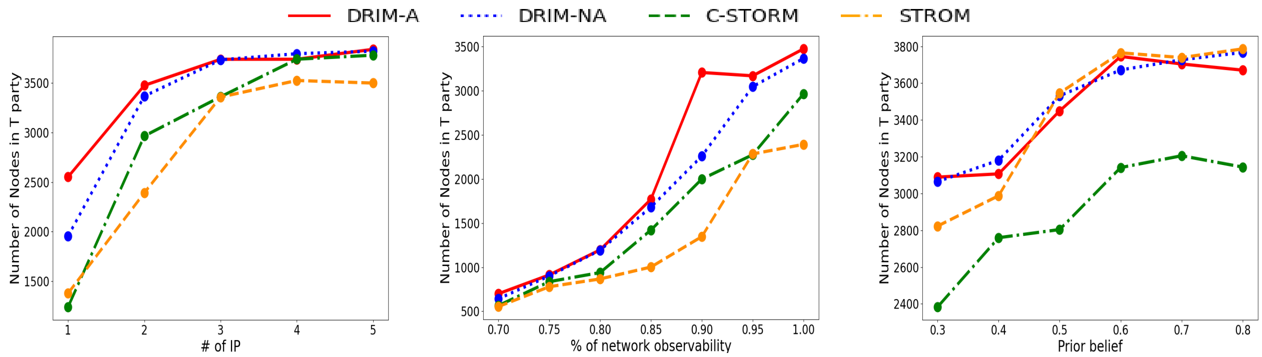
ACKNOWLEDGEMENT

This work is partly supported by NSF under Grant Numbers 2107449, 2107450, and 2107451.

REFERENCES

[1] K. Ali, C. Wang, M. Yeh, and Y. Chen. 2020. Addressing Competitive Influence Maximization on Unknown Social Network with Deep Reinforcement Learning. In *2020 IEEE/ACM Intl. Conf. on Advances in Social Networks Analysis and Mining (ASONAM)*. 196–203.

[2] K. Ali, C. Wang, M. Yeh, C. Li, and Y. Chen. 2021. NEDRL-CIM:Network Embedding Meets Deep Reinforcement Learning to Tackle Competitive Influence Maximization on Evolving Social



(a) Varying # of IP in FB Net.

(b) Varying % of NO in FB Net.

(c) Varying users' PB in FB Net.

Fig. 3: TP's influence across CIM algorithms when TP and FP use DRL for seed selection in the FB network. #IP denotes the number of information propagation by TP, 'NO' represents network observability %, and PB indicates users' prior belief in true information (i.e., a). In all cases except (c), we set $a = 0.5$.

Networks. In *2021 IEEE 8th Intl. Conf. on Data Science and Advanced Analytics (DSAA)*. 1–9.

[3] S. Bharathi, D. Kempe, and M. Salek. 2007. Competitive Influence Maximization in Social Networks. In *Internet and Network Economics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 306–311.

[4] J.H. Cho, S. Rager, J. O'Donovan, S. Adali, and B. Horne. 2019. Uncertainty-based False Information Propagation in Social Networks. *ACM Transactions on Social Computing* 2, 2 (2019), 1–34.

[5] T. Chung. 2019. Deep Reinforcement Learning-based Approach to Tackle Competitive Influence Maximization. In *Proc. MLG Workshop*.

[6] P. Domingos and M. Richardson. 2001. Mining the Network Value of Customers. In *Proc. the Seventh ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining*. 57–66.

[7] M. Granovetter. 1978. Threshold Models of Collective Behavior. *Amer. J. Sociology* 83, 6 (1978), 1420–1443.

[8] A. Jøsang. 2016. *Subjective Logic: A Formalism for Reasoning Under Uncertainty*. Springer.

[9] A. Jøsang, J.H. Cho, and F. Chen. 2018. Uncertainty Characteristics of Subjective Opinions. In *21st FUSION*. 1998–2005.

[10] V. Karnowski, L. Leonhard, and A. Kümpel. 2018. Why Users Share the News: A Theory of Reasoned Action-based Study on the Antecedents of News-sharing Behavior. *Communication Research Reports* 35, 2 (2018), 91–100.

[11] D. Kempe, J. Kleinberg, and É. Tardos. 2003. Maximizing the Spread of Influence through a Social Network. In *Proc. 9th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining (KDD '03)*. Association for Computing Machinery, New York, NY, USA, 137–146.

[12] D. Kempe, J. Kleinberg, and É. Tardos. 2003. Maximizing the Spread of Influence Through a Social Network. In *Proc. 9th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining*. 137–146.

[13] J. Leskovec and J. Mcauley. 2012. Learning to discover social circles in ego networks. *Advances in neural information processing systems* 25 (2012).

[14] J. Li, T. Cai, K. Deng, X. Wang, T. Sellis, and F. Xia. 2020. Community-diversified Influence Maximization in Social Networks. *Information Systems* 92 (2020), 101522.

[15] S. Lin, S. Lin, and M. Chen. 2015. A Learning-Based Framework to Handle Multi-Round Multi-Party Influence Maximization on Social Networks. In *Proc. the 21th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining*. New York, NY, USA, 695–704.

[16] W. Liu, L. Chen, X. Chen, and B. Chen. 2020. An Algorithm for Influence Maximization in Competitive Social Networks with Unwanted Users. *Applied Intelligence* 50, 2 (2020), 417–437.

[17] K. Michael, L. Markus, R. Mario, and S. Markus. 2021. Benders decomposition for competitive influence maximization in (social) networks. *Omega* 100 (2021), 102264.

[18] M. Nasim, R. Charbey, C. Prieur, and U. Brandes. 2016. Investigating Link Inference in Partially Observable Networks: Friendship Ties and Interaction. *IEEE Transactions on Computational Social Systems* 3, 3 (2016), 113–119.

[19] C. Pham et al. 2019. Competitive Influence Maximization within Time and Budget Constraints in Online Social Networks: An Algorithmic Approach. *Applied Sciences* 9, 11 (2019).

[20] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov. 2017. Proximal Policy Optimization Algorithms. *arXiv preprint arXiv:1707.06347* (2017).