

# Misinformation Propagation in Online Social Networks: Game Theoretic and Reinforcement Learning Approaches

Tolga Yilmaz<sup>ID</sup> and Özgür Ulusoy<sup>ID</sup>, *Member, IEEE*

**Abstract**—Misinformation in online social networks (OSNs) has been an ongoing problem, and it has been studied heavily over recent years. In this article, we use gamification to tackle misinformation propagation in OSNs. First, we construct a game based on the notion of cooperative games on graphs where the nodes of the social network are players. We use random regular networks and real networks in our simulations to show that the constructed game follows evolutionary dynamics and that the outcome of the game depends on the relation between the structural properties of the network and the benefit and cost variables defined in a cooperative game. Second, we create a game on the network level where the players control a set of nodes. We define agents whose goal is to maximize the total reward that we set up to be the number of nodes affected at the end of the game. We propose a deep reinforcement learning (RL) technique based on the multiagent deep deterministic policy gradient (MADDPG) algorithm. We test the proposed method along with well-known node selection algorithms and obtain promising results on different social networks.

**Index Terms**—Cooperative games, misinformation propagation, online social networks (OSNs), reinforcement learning (RL).

## I. INTRODUCTION

WHERE people receive news now includes online media such as news websites and social networks in addition to traditional media such as TV, radio, and newspapers [1], mainly due to the convenience in terms of quickness and socializing aspects [2]. These platforms have become sources to spread not just news but also ideas and have been utilized as tools to influence people for different purposes such as altering the public mind to vote for certain parties [3], advertising certain products for commercial advantage [4] generating awareness for certain issues in health [5], and global problems [6], [7].

The fast progress in communication does not come without disadvantages. With the fast pace of information exchange, the validation process also seems to be done faster, less thoroughly, and completely overlooked in most cases [8]. People tend to believe the information they see on the Internet and even help it propagate to others on social networks [9].

It is evident from the literature that misinformation spread is encouraged by various threat actors with the help of fabricated

text, media, and an army of ingenuine accounts that are used in the process. There is an ongoing game maintained by various actors, most frequently for various reasons such as political propaganda that could have profound effects on both internal and international stages.

Most of the time, the spread of misinformation does not occur as an isolated incident but in a repeated fashion. The users of various social media platforms encounter various stages and forms of misinformation daily, and so far, there does not seem a permanent solution. Therefore, it might be useful to address the misinformation problem as a repeated game in a social network environment where there exist multiple stages with multiple actors that affect the end-user (whether they participate or not) in some way. In this article, we define the misinformation problem as a game from two different perspectives: 1) node level and 2) network level.

In the first part of this work, we approach the misinformation propagation problem from a game-theoretic perspective. The setting where spatial relation between players plays a role in determining the outcome of the game is called spatial games [10], [11]. In the particular context of social networks as graphs, we approach the problem as a graph game. A vast amount of literature exists on graph games, and a particular study [12] approaches the problem of cooperation on graphs. Their work shows that cooperation is only favorable if players' benefit/cost ratio exceeds the average number of degrees (i.e., connections) in the graph. They show that otherwise, cooperation is not favorable. We construct a misinformation exchange game based on cooperative game theory. Our simulations also yield that the probability of countering misinformation is increased if the benefit/cost ratio exceeds the average degree. This approach, however, displays some disadvantages, such as the ambiguity of figuring out the actual or expected values of benefits and costs.

We then build another game that is played at the network level, where players are given a set of nodes and try to maximize the number of affected nodes. Since we find this setup suitable for a learning environment for a multiagent reinforcement learning (RL) setting, we propose to utilize a method based on the multiagent deep deterministic policy gradient (MADDPG) algorithm. For the same setting, we also implement other algorithms based on the highly established centrality, page-rank, and cost effective lazy forward (CELF) algorithms and share the results of the methods as a comparison.

Manuscript received 17 March 2022; revised 16 August 2022; accepted 19 September 2022. Date of publication 30 September 2022; date of current version 6 December 2023. (*Corresponding author: Tolga Yilmaz.*)

The authors are with the Department of Computer Engineering, Bilkent University, 06800 Ankara, Turkey (e-mail: tolga.yilmaz@bilkent.edu.tr).

Digital Object Identifier 10.1109/TCSS.2022.3208793

2329-924X © 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.  
See <https://www.ieee.org/publications/rights/index.html> for more information.

TABLE I  
PAYOFF MATRIX OF PRISONER'S DILEMMA

	Cooperate	Defect
Cooperate	$q = (-1, -1)$	$r = (-3, 0)$
Defect	$s = (0, -3)$	$t = (-2, -2)$

TABLE II  
PAYOFF MATRIX OF THE COOPERATION GAME

	Cooperate	Defect
Cooperate	$(b - c)$	$(-c)$
Defect	$(b)$	$(0)$

To summarize, our contributions are as follows.

- 1) We explore the misinformation propagation problem as a cooperative game on graphs and construct the payoff matrix where we define the benefit and costs associated with certain actions of the players (the nodes).
- 2) We run various simulations on random regular networks and a real network based on Facebook to explore how the game evolves concerning the benefit and cost ratio.
- 3) We define another game at the network level where there are agents that try to maximize the number of nodes affected on their behalf during their misinformation and counter-misinformation campaigns.
- 4) We propose a deep RL-based method for the selections of the actions of the agent in the network level game, perform experiments and compare the results with other node selection techniques.

## II. MISINFORMATION GAME AT THE NODE LEVEL

### A. Evolution of Cooperation on Graphs

Prisoner's dilemma is a well-known game in the field of game theory with two accomplices presented with two strategies: 1) cooperate (silence) and 2) defect (betrayal) [10]. A typical payoff matrix for the cooperation game inspired by the prisoner's dilemma is shown in Table I. In this scenario, although the mutual cooperation  $q = (-1, -1)$  is the most favorable overall outcome, the individually best choices force the players to defect:  $t = (-2, -2)$ .

Cooperation within communities has also been studied [13], including the iterated form of the game in a networked scenario to identify the evolutionary properties of cooperation [14]. In such a game, people are nodes on a graph and connected with edges representing a contextual relationship. A cooperator emits a benefit  $b$  to all its neighbors at the cost of  $c$ . A defector does not emit any benefit but benefits from the cooperators it neighbors. The payoff matrix of such a game is given in Table II.

In this networked setting, with benefit and cost presented as the game parameters, the evolution mechanism is layout through updating nodes with respect to certain mechanisms. At a time instant, a node is randomly chosen to be updated. There are three such update strategies [12].

- 1) Death-birth updating strategy where a node is chosen to die and cooperators and defectors compete over it without considering its original fitness.

- 2) Birth-death updating where a node is chosen for reproduction and its child replaces one of the neighbors.
- 3) Imitation updating where a node is chosen to replace its strategy with respect to its neighbors. Its fitness is taken into account.

We will specifically visit death-birth updating and imitation updating strategies since these two can provide intuitive applications in today's online social networks (OSNs): death-birth updating can simulate setting the strategy of a node only by its neighboring strategies, while imitation updating can provide an analogy for the update of a node while also considering its own fitness in the final decision.

Here, we will describe the strategies using the payoffs in Table I. Under death-birth updating, let us assume that a node is to be updated. Let  $k$  represent the number of neighbors of the node, a  $C$  player is a cooperator, and a  $D$  is a defector. In this sense,  $k_C$  represents the number of  $C$  neighbors whereas  $k_D$  represents the number of  $D$  neighbors with  $k = k_C + k_D$ . Then, the node's strategy is set as  $C$  based on the probability as follows:

$$\frac{k_C f_C}{k_C f_C + k_D f_D} \quad (1)$$

where  $k_C f_C$  and  $k_D f_D$  denote the total fitness of the neighboring nodes with  $C$  and  $D$  strategies, respectively. The fitness of a  $C$  player and a  $D$  player is described in the following equations if the selected node was a  $D$  player:

$$f_C = 1 - w + w[(k-1)p_{C|C}q + [(k-1)p_{D|C} + 1]r] \quad (2)$$

$$f_D = 1 - w + w[(k-1)p_{C|D}s + [(k-1)p_{D|D} + 1]t]. \quad (3)$$

In these two equations, we describe the fitness of a node in terms of its neighbors.  $p_{C|C}$  is the probability of finding a  $C$  node as a neighbor of a  $C$  node, whereas  $p_{D|C}$  is the probability of finding a  $D$  node as a neighbor of a  $C$  node.  $(k-1)p_{C|C}q$  represents the total contribution from  $C$  neighbors. Calculating these probabilities is relatively cost-effective in large graphs.  $w$  is the selection parameter that provides a linear combination within the game dynamics. We talk about a strong selection when  $w$  is close to 1; a weak selection when it is very small.

For imitation updating, we need to describe the fitness of a  $C$  node  $f_0$  that is about to be updated as follows:

$$f_0 = 1 - w + w(k_C s + k_D t). \quad (4)$$

The node chooses strategy  $D$  with respect to the following probability:

$$\frac{k_D f_D}{k_C f_C + k_D f_D + f_0}. \quad (5)$$

According to the game plan where payoffs are set as in Table II, cooperators are favored if  $b/c > k+2$  under imitation updating and  $b/c > k$  under death-birth updating [12].

### B. Information Propagation as an Evolutionary Game on Graphs

We visit similar games from the literature toward an analogy for the misinformation game. "Closed-bag exchange" is a

game where two people exchange bags containing money and goods. In the game, players either honor the deal (cooperate) or deliver an empty bag (defect). “Peace war game” is another example where making peace (cooperate) is mutually beneficial; the one-sided “war” (defect) strategy brings more benefit to the game. Finally, the lying game has been modeled in numerous research [15]. Strategizing for the interest of the individual—as evident in these games—may lead to a notion called the tragedy of the commons for shared resources. The problem has been translated to the digital world as the tragedy of the digital commons, and the lack of regulatory systems causes pollution in digital resources also associated with misinformation [16], [17].

In an OSN, some nodes may be inclined or even work for misinforming their proximity (i.e., their followers, connections, friends) on purpose. This deliberate version of misinformation is called disinformation, and the nodes are responsible for their actions and thus can be included in a game-theoretic environment. After being misinformed, previously indifferent nodes can relay this information to other nodes, making them a part of the game. There may exist other nodes that work on the opposite side of the disinformers. These nodes have to work harder than the latter since it is harder to convince the other nodes about the truth, while false information is generally more catchy and sticky, or interesting. The research confirms that false information spreads faster [18], [19].

First, let us define an information exchange game between two parties. Assume there are two strategies: 1) cooperation and 2) defection, where cooperation is sending correct information while defection is sending a false one. A player receives benefit  $b$  if the other player chooses to share correct information and there is a cost of correct information to the sender, while false information provides no benefit to the receiver. The payoffs are defined exactly like the cooperation game defined in Table II. The Nash equilibrium of this game is with the outcome  $(0,0)$  where both players choose to disinform. However, there was a better outcome for them  $(b - c, b - c)$  if they could both choose to relay the correct information. This game can also be intuitively connected to the famous closed-bag exchange game. If the game is set up around this description, the studies about the evolution of cooperation on graphs can be easily applied and the findings are expected to be in parallel. Although this type of setting for the game enables observations on the adaptation of cooperative or defective strategies over the population, it does not closely simulate the properties of propagation through iterative rounds of games over time. Hence, further modifications are required.

In this modified game, the possible strategies for player 2 are different. Player 1, again, shares either correct information or a false one while player 2 either accepts the information as correct, i.e., believes it, or does not accept it. Since this is an iterated game, player 2 can then become a spreader in the later rounds. In this new setting, it is possible to define fine-grained values for benefit and cost for each player. The utility of the correct information to the sender is  $b_1$  with a cost of  $c_1$ . The utility of receiving correct information to the receiver is  $b_2$  if the receiver believes and  $b_3$  if the receiver does not believe. That is because there is an intrinsic value in

TABLE III  
PAYOFF MATRIX OF THE MISINFORMATION GAME

	Believe	Don't Believe
Correct	$(b_1 - c_1, b_2) \approx (b - c, b - c)$	$(-c_1, b_3) \approx (-c, b)$
False	$(b_4 - c_2, -c_3) \approx (b, -c)$	$(-c_2, 0) \approx (0, 0)$

the correct information. However, the sender does not receive any benefit for the latter. The utility of the false information to the sender is  $b_4$  if the receiver believes it, and 0 otherwise. The cost of false information is  $c_2$ . Believing false information has a cost of  $c_3$ . Although the information exchange game we previously introduced had the same payoff matrix, we need to introduce some assumptions for the misinformation game in its current form. First, we assume  $c_1$  and  $c_3$  as equal and denote it as  $c$ . We assume that  $c_2 \ll c_1$  and disregard  $c_2$ . We also take all benefits values as equal, except for  $b_2$ . We think that it is upper-bounded by  $b - c$  since the utility of a receiver cannot be larger than the sender's if the sent information is correct. The simplified payoff matrix is given in Table III. The Nash equilibrium of this simplified game is  $(0,0)$ , to disseminate “False” information for player 1 and “Do not Believe” the information for player 2. However, the scenario of sending “Correct” information and “Believing” provides a better and mutually beneficial outcome for players 1 and 2, respectively. Hence, the resulting non-zero-sum game displays the same characteristics of the cooperative games and the prisoner's dilemma in particular, under the previously listed assumptions on the benefit and cost values.

In addition, while the Nash equilibrium provides a general solution for games in the traditional setup, we may need other measures of evolutionary dominance under evolutionary settings. Evolutionarily stable strategy (ESS) [20] is a modification of the Nash equilibrium, which states that a strategy is said to be evolutionarily stable if adopted by a population in an evolutionary environment, and cannot be replaced by another strategy. Given  $E(I, J)$  as the payoff of selecting strategy  $I$  against  $T$ , for the strategy  $I$  to be an ESS, two conditions should be considered [20] as follows:

- 1)  $E(I, I) > E(J, I)$ .
- 2)  $E(I, I) = E(J, I)$  and  $E(I, J) > E(J, J)$ .

According to this definition, the defection strategy is evolutionarily stable in the designed misinformation game. However, it has been shown that the evolution of cooperation is possible in the case of  $b/c > k$  [12] and small  $k$  or large  $w$  are the two factors that affect the outcome in favor of cooperation in the iterated cooperation game on graphs [21].

1) *Combined Strategy for the Misinformation Game*: When the misinformation game is played out on the network, there exist three types of actors: 1) defectors; 2) cooperators; and 3) neutral nodes. In a social network, these correspond to the misinformers, correctors, and neutral nodes (red, blue, and gray nodes, respectively). In epidemiology as well as information diffusion theory, there exist two main types of notions that describe the state of nodes in a network: 1) susceptible, infected, and recovered model (SIR) (Kermack and McKendrick in 1927); and 2) susceptible,

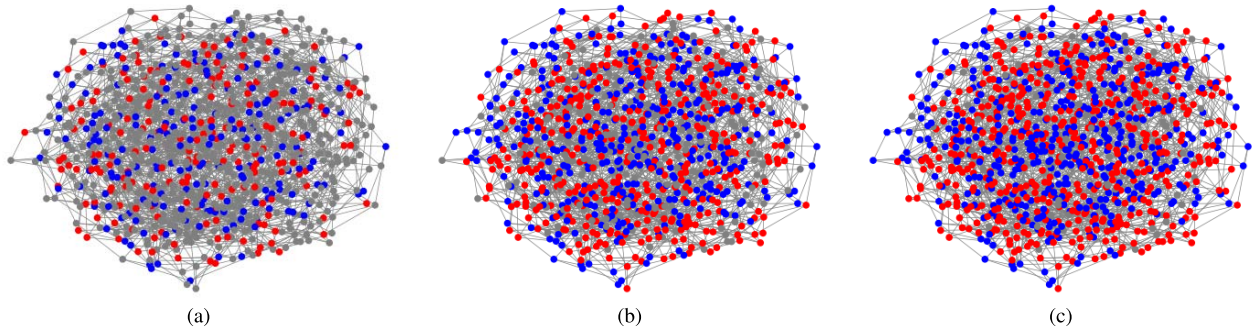


Fig. 1. Random network experiment 1:  $k = 4$  and  $b/c = 2$ . (a) Initial setting. After (b) 2500 iterations and (c) 5000 iterations.

---

### Algorithm 1 Node Update Algorithm

---

**Require:**  $G$ : Graph

**Require:**  $B$ : Set of Cooperator Nodes (Blue)

**Require:**  $R$ : Set of Defector Nodes (Red)

**Require:**  $Gr$ : Set of Neutral Nodes (Gray)

```

1: procedure UPDATENODE( $G, B, R, Gr$ )
2:    $S_B \leftarrow 0$  ▷ Blue Fitness Sum
3:    $S_R \leftarrow 0$  ▷ Red Fitness Sum
4:    $v, V \leftarrow \text{randomlychoose}(G, B, R, Gr)$ 
5:   for  $\hat{v} \in V$  do
6:      $f_{\hat{v}} \leftarrow \text{calculatefitness}(G, \hat{v}, B, R)$ 
7:     if  $\hat{v} \in B$  then
8:        $S_B \leftarrow S_B + f_{\hat{v}}$ 
9:     else if  $\hat{v} \in R$  then
10:       $S_R \leftarrow S_R + f_{\hat{v}}$ 
11:     end if
12:   end for
13:    $f_v \leftarrow \text{calculatefitness}(G, v, B, R)$ 
14:   if  $v \in B$  then
15:      $S_B \leftarrow S_B + f_v$ 
16:   else if  $v \in R$  then
17:      $S_R \leftarrow S_R + f_v$ 
18:   end if
19:   if  $S_B > S_R$  then
20:      $B \leftarrow B \cup v$ 
21:   else if  $S_B < S_R$  then
22:      $R \leftarrow R \cup v$ 
23:   end if
24:   return  $v$ 
25: end procedure

```

---

infected, susceptible (SIS) model. In the SIR model, a node can be susceptible, infected, or recovered without ever getting infected again. In the SIS model, however, a node can be reinfected. With this analogy, a susceptible, hence neutral node, can be infected or misinformed and become a spreader. It can be recovered with correct information to become a corrector. In this work, we chose the SIS model; thus, it is possible that a node can change its state from a gray node to a red or blue node, and a blue or red node can invert its position to become a red or blue node. To reflect this strategy, we need to accommodate two different strategies: 1) one for

gray nodes to select a new position and 2) one for nodes with an existing stance to change their type. The first corresponds to a death–birth updating strategy where the fitness of the node to be updated is not taken into consideration. The latter is when a node updates its type based on its neighbors and its own fitness.

We adopt the updating algorithm given as Algorithm 1. This algorithm denotes a mix of death–birth and imitation updating strategies. The algorithm is described as follows. During the simulation, a node  $v$  is randomly selected to be updated, along with its set of neighbors  $V$ . For each neighbor  $\hat{v} \in V$ , a fitness value  $f_{\hat{v}}$  is calculated and it is added to a cumulative sum;  $S_B$  for cooperators (blue) or  $S_R$  for defectors (red), according to the strategy (blue or red) of  $\hat{v}$ . Then, the fitness of the selected node  $f_v$  is calculated. After the addition of  $f_v$  to the cumulative sum of its original strategy, the strategy of  $v$  is updated with strategy  $B$  or  $R$  with the larger cumulative sum ( $S_B$  or  $S_R$ ).

### C. Simulations

In this section, to observe whether the misinformation game described in this work is similar to the cooperation game, we run various simulations. The particular point we are after is the inequality of  $b/c > k$ . We want to see whether the graph is to be dominated by misinformation when the inequality fails, and whether correct information holds when the inequality is met.

1) *Random Regular Networks*: To test with the changing number of average neighbors, we choose to experiment with random regular graphs. Furthermore, we apply a community detection algorithm to create groups of nodes. These will serve as the set of competing nodes over unassigned nodes. In all our experiments, red nodes describe misinformers, and blue nodes describe correctors. Gray nodes are not assigned. We also pay attention to the sizes of the blue and red groups. We do not want one group to have a larger upstart advantage over the other.

In the first experiment,  $k$  is chosen as 4 and  $b/c$  is 2. Fig. 1 shows the state of the network, and we see that the misinformers strategy increases its population over the cooperator strategy. In the second experiment,  $k$  is chosen as 4 and  $b/c$  is 20. We see that blue nodes end up with a slightly larger population than the red nodes (Fig. 2).

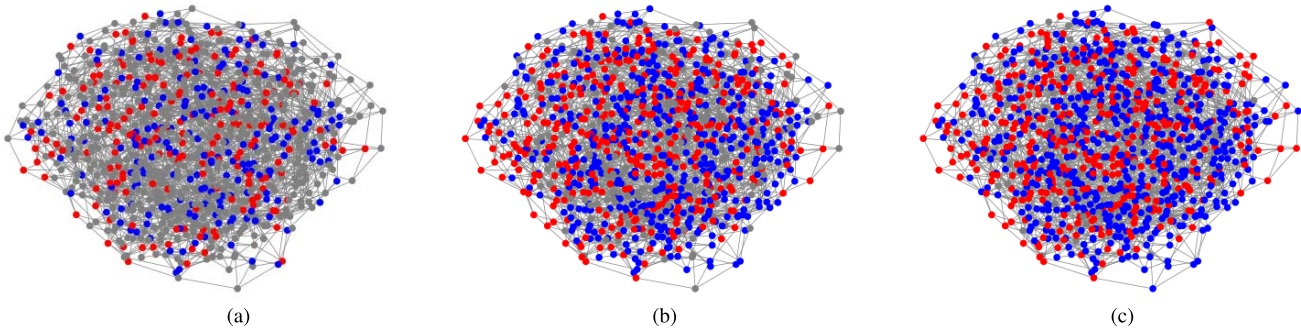


Fig. 2. Random network experiment 2:  $k = 4$  and  $b/c = 20$ . (a) Initial setting. After (b) 2500 iterations and (c) 5000 iterations.

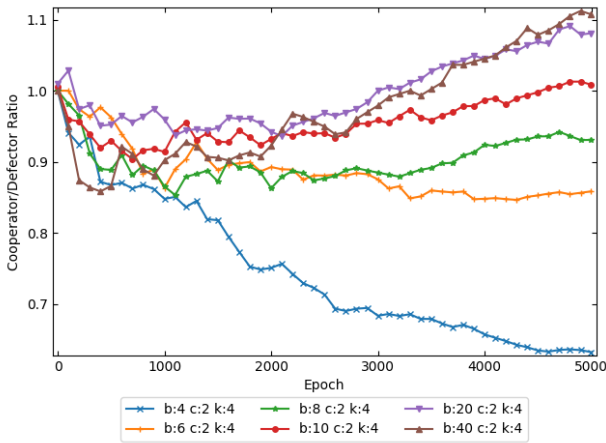


Fig. 3. Change of blue/red ratio during the simulation of random network experiment with various benefit and cost values.

Fig. 3 shows the change of blue/red ratio over time for different values of benefit and cost when  $k = 4$ . According to the experiments, the dominance of cooperators is possible if  $b/c > k$ .

2) *Facebook Network*: We also run simulations on a real data set based on a set of Facebook users published in [22] (around 4000 nodes, 80 000 edges). Similarly, we first choose two communities and label these as blue or red. Then we run the same algorithm. In this graph  $k = 43$ . In the first experiment with this data, we choose  $b/c = 5$ . The results in Fig. 4 show that there is red dominance.

In the second experiment, we exaggerate the ratio of  $b/c$  to see its effect. Fig. 5 shows that the blue strategy dominates the red strategy with this setting.

We repeat the experiment for various  $b/c$  combinations, and again, according to our observations, as shown in Fig. 6, the cooperators are favored if  $b/c > k$ .

3) *Strategy Dominance Probabilities*: In evolution and evolutionary game theory, as well as in [12], the term fixation corresponds to the state of a network where a network is completely covered in one of the types of nodes. In our simulations, at this moment of our research, we use the “win” probability, which we obtain by the ratio of blue wins over red over a finite number of iterations. In Fig. 7(a), we show

that the winning probability increases as  $b/c$  increases. The graph shows data for small ( $N = 100$ ), medium ( $N = 1000$ ), and large ( $N = 10\,000$ ) networks, each with  $5N$  epochs and 50 iterations.  $k$  was chosen as 4.

4) *Effect of Initial Node Distribution on the Dominance*: In our previous simulations, the initial network setting was the random distribution of nodes. However, in real-world scenarios, this may not be the case. In Fig. 7(b), we start the network after calculating two same-sized clusters for opposing sides using community detection where most of the network is neutral. We use the Leuven method [23] for community detection for its wide acceptance and accessibility. However, more recent methods such as [24] and [25] could also be used. Our initial results show that the win probability is dramatically increased in this network structure.

In the community setting, the connectivity between the clusters is low (local  $k$  is low) at the very beginning. Since red nodes require blue nodes to benefit, initially, the spread rate of red nodes is low; only when the connectivity between red nodes and blue nodes is high (local  $k$  is high) then the red nodes are advantageous.

D. Limitations of a Node-Level Game

Studying the misinformation game where the players are the nodes within a social network may enable a better theoretical understanding of how nodes change their strategies under an evolutionary setting. On the other hand, establishing such a game in contemporary social networks on the web in a holistic manner is difficult. This is because the users are actual people with different aspirations and have different motives for using such networks, and they are exchanging and influenced by different information simultaneously. There may be individual benefit and cost values for each interaction rather than static network-wide values for them. In addition, determining benefit and cost as discrete variables that could simulate or offer analogies for the real-life benefits and costs of the said misinformation mechanism is also difficult; hence as was in our study, it leads to making assumptions.

However, it may be beneficial to list some of the concepts we considered for the values of benefit and cost while doing the study and their limitations. One such example would be to associate benefit with reputation. In this context, sending correct information would yield some benefit in the form of

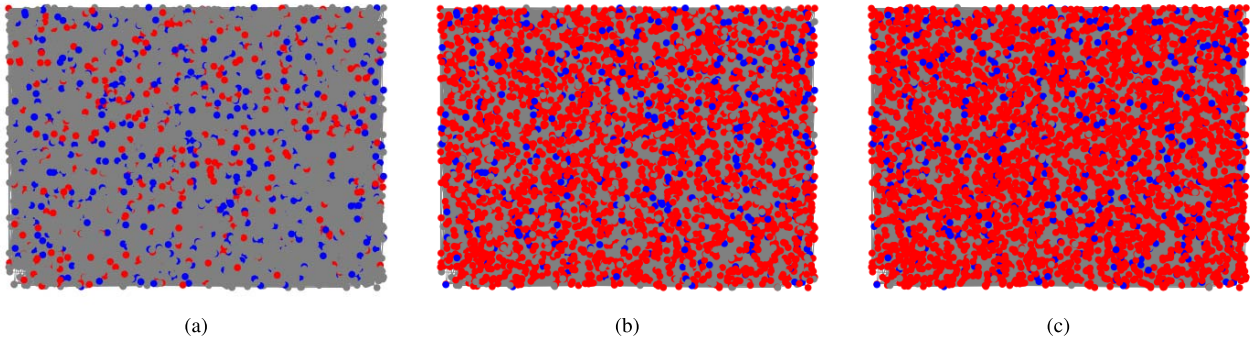


Fig. 4. Facebook experiment:  $k = 43$  and  $b/c = 5$ . (a) Initial setting. After (b) 10000 iterations and (c) 20000 iterations.

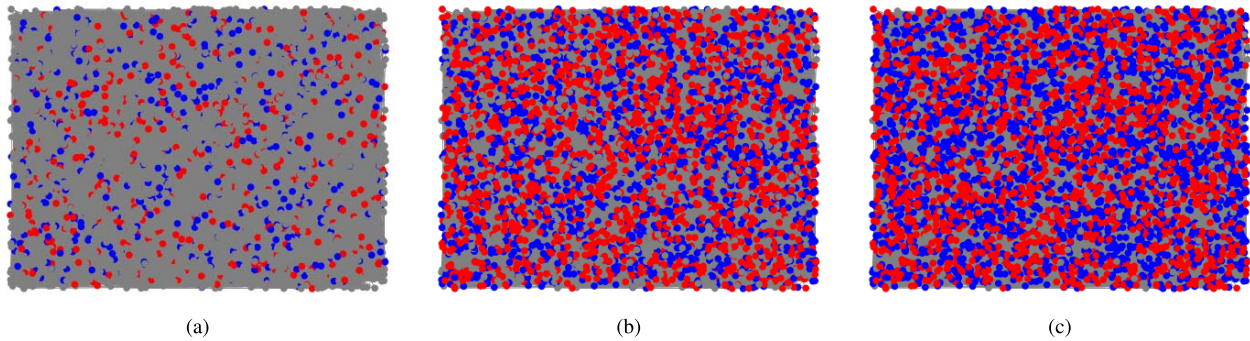


Fig. 5. Facebook experiment  $k = 43$  and  $b/c = 70$ . (a) Initial setting. After (b) 10000 iterations and (c) 20000 iterations.

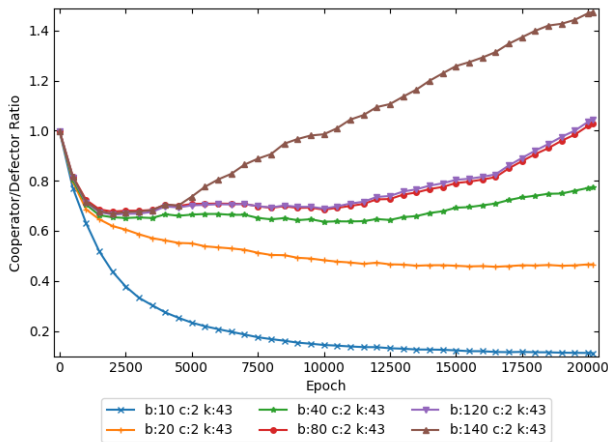


Fig. 6. Change of blue/red ratio during the simulation of Facebook network with various benefit and cost values.

reputation, while cost means preparing such information. The problem with this is that it is not intuitive to represent the value of information in terms of reputation, and vice versa, so that we can calculate the payoff, not to mention the hardness of deciding what reputation is. It may be possible to associate the said cost with the expected loss of reputation if we ignore the value of information. Yet, it may be possible to incorporate the value of information into the expected reputation. However, we would still need to differentiate between correct and false information.

We showed that the average number of neighbors  $k$  is indeed a factor in how the evolution of strategies among nodes occurs. In addition, we showed through simulations that prior predispositions such as existing communities (e.g., cliques or clusters) would also indicate different evolution characteristics. Prior studies on spatial evolutionary games also show such results regarding the effects of network structure [26], [27]. This motivates future work on a fine-grain analysis of the effects of connectivity, such as the size and the number of cliques, echo chambers, and types of relationships. For instance, if the value of benefit and cost were dynamic, as previously said, the feasibility of a partitioning algorithm based on the node-wise values of  $b/c > k$  could be studied.

### III. GAME BETWEEN NETWORK-LEVEL PLAYERS

In today's OSNs, there exist intrinsic actors above the node level, i.e., outside the network, with different motives such as politics, advertisement, and reputation who try to spread various information to affect people's minds using various techniques. One such technique is to control or influence a set of nodes that serve for the benefit of the actor during an information spread campaign. These nodes can be maintained by real people (sometimes called trolls) or could be bot accounts [28]. A vast amount of research exists on identifying and mitigating fake accounts in OSNs, and a recent review is provided by [29]. While dealing with misinformation through means of identifying such ingenuine accounts provides relief for the real people to be notified about such accounts and help regulate the social network, the broader problem specification

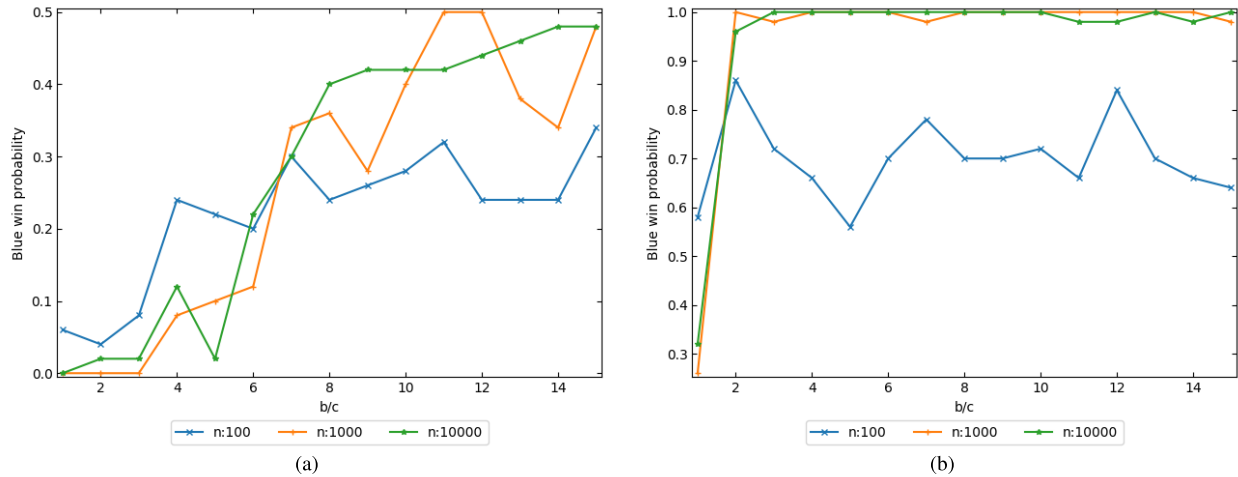


Fig. 7. Strategy dominance probabilities with different network sizes and  $b/c$ . Blue win probability with changing (a)  $b/c$  and (b)  $b/c$ , network partitioned using community detection.

is mainly associated with the area of influence maximization which deals with identifying the parameters that lead to maximal influence for various agents over the nodes of social networks. Influence maximization has been identified as an NP-hard problem [30].

Given a social network, which is a graph with nodes representing users and edges (directed or undirected) representing a relationship (such as friendship, follow, connection), there is at least one player that controls or influences directly some of the nodes to start spreading some information. The scenario becomes misinformation propagation if the information spread falls into the misinformation category. The purpose of the player is to maximize the number of nodes affected.

#### A. Mechanism Design

We set the environment for the game to be the network. For the sake of simplicity, there exist two players, each given a set of randomly selected nodes. While player one spreads misinformation, the other player opposes the misinformation campaign. In each time step, each of the players utilizes one of the nodes as the seed for misinformation. As the propagation mechanism, we chose the SIR model. The reward for the players is the number of affected nodes after the game is ended. In this work, we utilize various well-known node selection algorithms and propose another one based on deep RL using the MADDPG algorithm. The nodes that are selected by the algorithms out of a randomly selected (same for each) set of nodes then are used in the information propagation game. Below, we describe the specifics of the baseline algorithms and the proposed method.

1) *Node-Centrality*: Centrality is a measure of a node's location in the network and is generally used to identify the importance of the node. There are various techniques for calculating the value, such as the number of in-degrees and out-degrees, eigenvector centrality, Katz-centrality, and others. In this work, we experimented with various node-centrality measures and opted for the degree-centrality method as we

observed that the results of those measures appear to be quite similar.

2) *Page-Rank*: Page-rank was introduced by the Google search engine to find out the importance of web pages. Today it has been modified and used in many areas, including social network analysis as a measure of node importance.

3) *Greedy*: The greedy algorithm was proposed by [30]. It takes a network with  $n$  nodes and computes the spread value until it finds  $k$  nodes with maximal marginal spread. Its complexity is  $O(kn)$  multiplied by the time required by the spread. Theoretical guarantees exist, mentioning that the algorithm achieves at least 63% of the spread resulting from the optimal set.

4) *Cost Effective Lazy Forward (CELF)*: CELF [31] is a modification of the greedy algorithm that achieves the same results with less computation using an optimization technique called lazy-forwarding.

5) *Proposed Method Based on MADDPG*: We approach the selection of nodes for misinformation or countering it from an RL perspective. The problem statement is as follows: the social network is an environment consisting of states, actions, and rewards. At any point in time, the state  $s$  is a list of node stances, the actions are a list of selected nodes, and the rewards are the number of nodes affected by the actions. Is it possible to learn a policy  $\pi$  that could maximize the expected reward over time? ( $\sum_t E[r_t | \pi]$ ) (see Fig. 8).

An RL problem can often be described as a Markov decision process (MDP), which contains the transition function that encapsulates the state-to-state transition probabilities and the reward function that outputs the value of the reward given the current state. In such a context, the transition and the reward functions can be thought of as the model of the environment and provide a basis for a subset of RL algorithms called the "model-based" algorithms which utilize the said model to find an optimal policy that gives the maximum expected reward.

However, in some cases, the definition, the transition probabilities, and the associated reward functions of an MDP are unknown for various reasons, such as the complexity of the

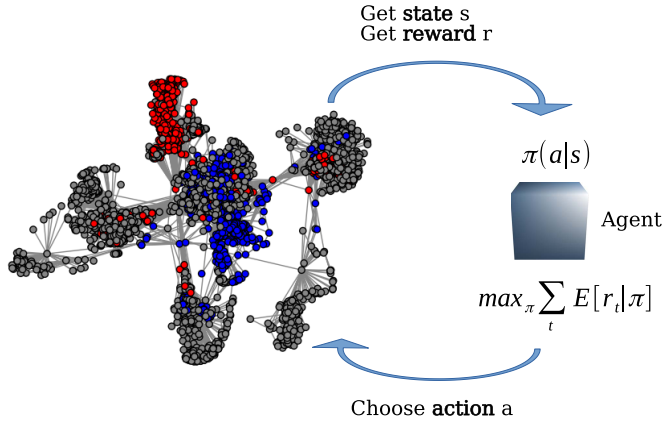


Fig. 8. RL problem over a social network.

environment or purely design choices. The RL algorithms that are specifically designed to learn in such environments are called the “model-free” RL algorithms. These do not utilize the transition and reward functions but rather often have a way to learn a value for the current state of the environment explicitly by interacting with the environment. This value function can then be used to determine a policy.

Q-learning can be recognized as the starting point of such approaches that are based on trial-error; however, as the action/observation spaces grow exponentially with respect to the complexity of the environments, the need for deep neural layers introduced other methods such as deep Q-network (DQN) instead of keeping track of every action-state tuple [32] in a Q-table. There are also policy gradient-based algorithms [33] which are used with continuous action spaces where a policy is a parametric distribution, and these parameters are adjusted using gradient descent. These algorithms led to actor-critic methods, deterministic policy gradient (DPG) algorithms [34], and an algorithm called deep DPG (DDPG) [35]. Two possible problems related to stability arise in the use of DPG algorithms. The first is related to the method being “on-policy”—which means that the critic evaluates the value of actions based on the same policy—creating possible bias [36], by disabling the utilization of a stabilization mechanism such as the experience replay buffer in DQN. The second issue is the sample complexity problem that is related to the required number of samples for efficient learning [37], [38]. In DDPG, there is a single agent with actor and critic networks where the actor-network chooses an action based on the state of the agent, and the critic network determines the value of that selection. To reduce the previously stated stability problems, DDPG first uses an experience replay buffer to store past transitions to operate “off-policy.” Second, it employs target networks associated with the actor and critic networks combined with a soft-update mechanism to increase stability. [39] iterates the possible failures and problems in DDPG that may result in poor learning.

MADDPG [40] was offered as an extension to DDPG for multiple agents. In MADDPG, all agents again have their actor

TABLE IV  
HYPER-PARAMETERS FOR THE MADDPG ARCHITECTURE

Hyper-parameter	Value	Hyper-parameter	Value
Actor Network	1024x512	Actor Optimizer	Adam
Critic Network	1024x512	Critic Optimizer	Adam
Actor Learning Rate $\alpha$	0.0005	Tau	0.01
Critic Learning Rate $\beta$	0.0005	Gamma	0.95
Batch Size	128		

TABLE V  
DATA SETS

Network	Type	Nodes	Edges	Reference
Facebook EgoNet	Undirected	4039	88234	[41]
Twitter EgoNet	Directed	81306	1768149	[41]
Facebook MIT	Undirected	6402	251230	[42]
Epinions	Directed	26588	100120	[43]

and critic networks; however, critic networks have full access to the environment. In addition, MADDPG utilizes a mechanism called the policy ensembles for more robustness, along with the inherited mechanisms from DDPG. Instead of relying on a single policy per agent, an ensemble of policies exists to sample from. In this work, we chose MADDPG as it reportedly outperformed various other methods [40] previously, and it supports continuous action spaces in multiagent environments. Also, the agents can see the actions of other agents (even if partially), which is suitable for the scenario in the scope of this work.

The MADDPG architecture is comprised of actor and critic networks along with their target networks. The environment is the set of  $n$  nodes, and the observations are the states of the nodes. Each node can have one of the three states, infected, neutral, or recovered (i.e., under-misinformation, neutral, or correctly informed). We set up the network to take the  $n$  node states as input, and the number of outputs is set out as the number of seed nodes  $s$ . The outputs are continuous values and are sorted at the end. The network chooses the output with the largest value in a sense. The other methods are also given the seed  $s$  nodes as the input and choose a subset of  $k$  nodes to be the originator nodes. This means that CELF, for instance, which is an algorithm that selects the best nodes in the network, is now modified to select the  $k$  best nodes from a subset of  $s$  nodes. This could potentially undermine the theoretical guarantees mentioned earlier. However, in real networks, the actors cannot choose nodes at will from the entire network but have to work with what they have. Nonetheless, the brute force algorithm still requires  ${}^s P_k = s!/(s-k)!$  number of cascades.

The hyper-parameters for the MADDPG architecture are given in Table IV.

## B. Experimental Results

In our experiments, we use four fairly large networks, two from Facebook, one from Twitter, and one from the Epinions.com dataset. The details are given in Table V. In our experiments, we randomly select  $s = 100$  nodes per game as the pool for selection for the algorithms. The algorithms



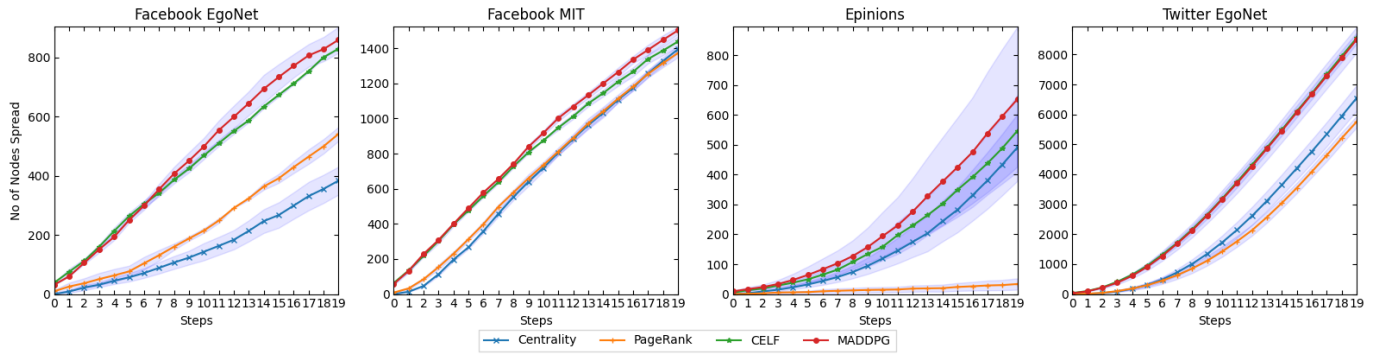


Fig. 9. Agent 1—cumulative spread.

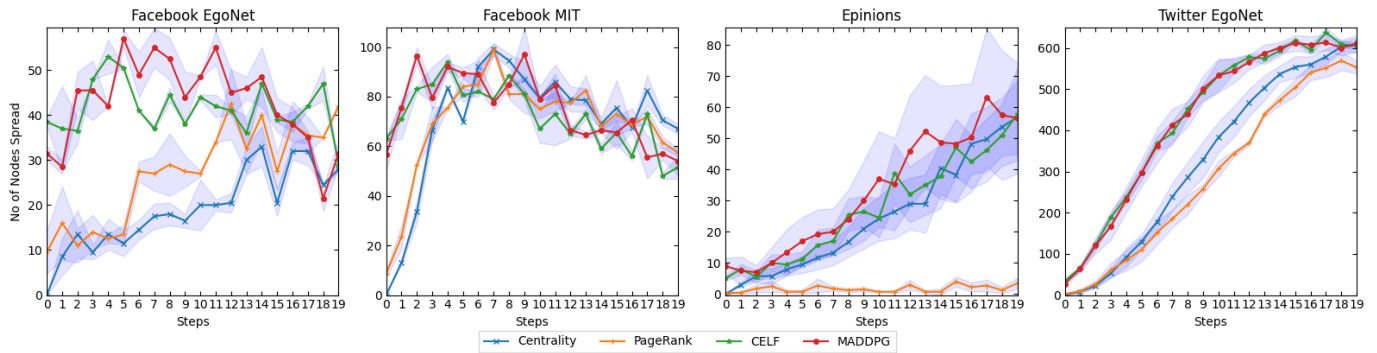


Fig. 10. Agent 1—spread at each step.

then select top  $k = 20$  nodes as originators. We then run one spread iteration per originator node and observe the total spread. We continue until all  $k$  nodes are exhausted. The spread dynamic is chosen as SIR, and we utilize the following transitions probabilities:  $P(I|S) = 0.02$ ,  $P(R|I) = 0.01$ ,  $P(R|S) = 0.01$ .

During the training, we incorporated two mechanisms for improving the agent networks. The first one is to introduce noise to facilitate learning. We found that the addition of noise is critical for exploration. While experimenting with various noise mechanisms, we decided on a noise function based on an Ornstein–Uhlenbeck process, also called the Vasicek model (6). The first part defines the drift over  $X$  where  $|X| = s$ .  $\mu$  defines the long-term mean, and  $\theta$  is the mean reversion speed. The second part  $dW_t$  is the discrete form of a Wiener process (7) at time step  $dt$ , where  $W$  is a random variable between  $[0, T]$  and  $\sigma$  is the scale of randomness, i.e., volatility. Given  $W_0 = 0$ ; for  $0 < s < t < u < v < T$ ,  $W_t - W_s$  and  $W_v - W_u$  are independent increments and these increments follow a Gaussian  $N$  distribution with zero mean and unit variance. The noise, then, is sampled at time step  $t$  and added as  $X_t + dX_t$

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t \tag{6}$$

$$dW_t \sim \sqrt{dt}N(0, 1). \tag{7}$$

The second improvement is to decide when to save checkpoints during the training by using a sliding window of size  $k$  for the past rewards. Here, there were many available options,

such as sum, mean, rolling sum, etc., but we used the area under the curve (AUC). If the window has a larger AUC than the previous best, we save the checkpoint.

We report the results of Agent 1—who tries to maximize the spread of misinformation, and Agent 2—who tries to minimize it. We experimented with various combinations of  $s$  and  $k$ , and as the results were similar, we only report the results in the mentioned setting. The MADDPG agents were trained a maximum of 1000 times per game. We played 100 games for the results. The results are given as the mean curves and the 95% confidence interval was also plotted. We omitted the results for the greedy algorithm as the results coincide with the CELF algorithm, as previously expected.

Fig. 9 contains the results for Agent 1 and the cumulative spread. The results show that the agent effectively learns the set of influential nodes as compared to other algorithms, and even outperforms an established algorithm—CELF in most cases. Fig. 10 gives the spread at each step. Figs. 11 and 12 give the cumulative and step-by-step spreads of Agent 2, respectively. The results are similar. We see that the improvement experienced by Agent 1 is also experienced by Agent 2. We see that the undirected networks show different characteristics than the directed ones. For directed graphs diffusion happens much faster considering the number of nodes spread. This may be due to the size difference between those networks and connectivity (e.g., the average number of neighbors) inside the network. We also notice that the spread amount for Agent 2 is around half of Agent 1 for directed graphs, which is expected as we set up the transition probabilities

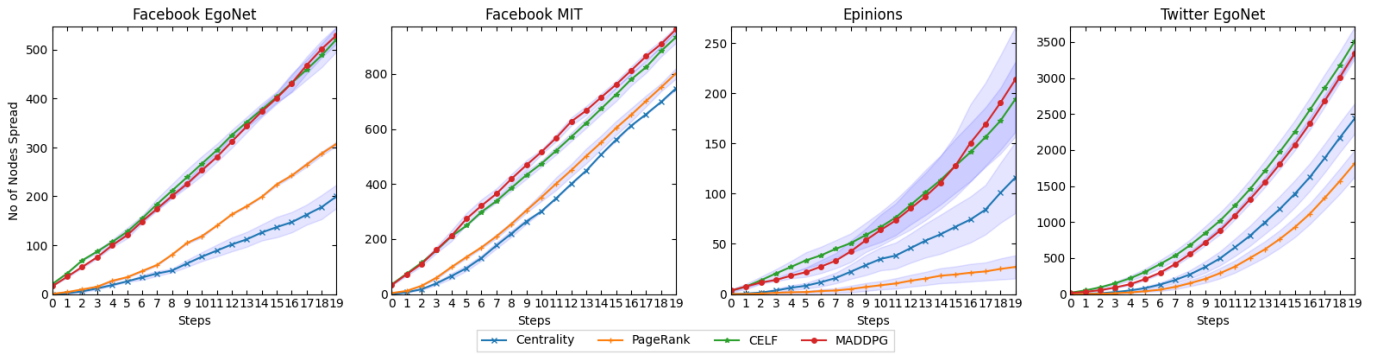


Fig. 11. Agent 2—cumulative spread.

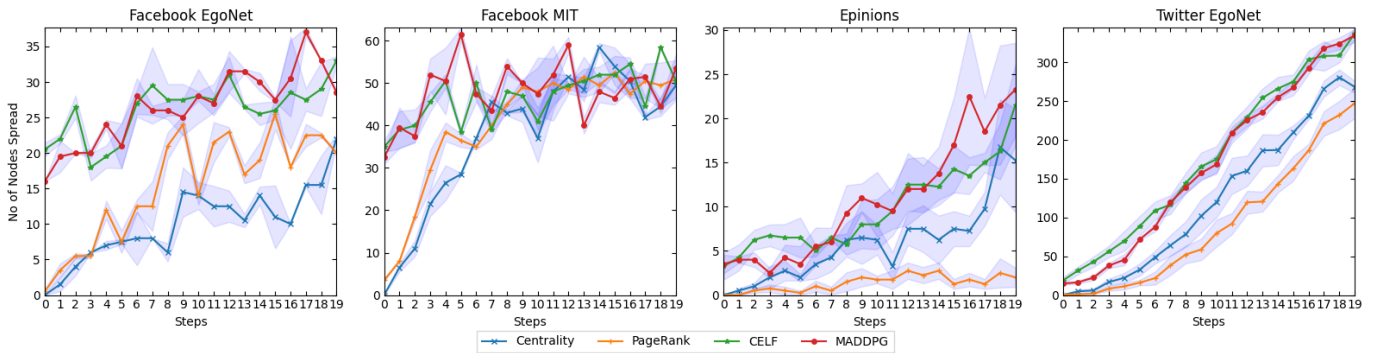


Fig. 12. Agent 2—spread at each step.

of the SIR model that way. However, Agent 2 seems less successful in directed graphs—Epinions and Twitter—than the undirected graphs, the cumulative spread not reaching half of Agent 1 for these networks. It should also be noted that the classical methods—centrality and page-rank—still seem practical choices for node selection tasks.

#### IV. DISCUSSION

One of the main issues of using a deep neural network is the interpretability of results, i.e., making sense of its choices. This also remains an issue in our work to be explored in the future.

In this work, we did not utilize any node representation techniques such as an adjacency matrix, convolutionary graphs nodes or a learned representation such as node2vec [44] or a network representation scheme such as averaging over node2vec embeddings, DeepWalk [45] or anonymous walks [46]. This situation creates two immediate limitations. First, it takes longer to train the network if we do not provide the node representations. Second, the trained agents cannot be generalized/transferred to work for other social networks but instead work for the trained network only. However, there are also opportunities in the approach. As the agents learn from the bare states of the nodes, the resulting actions could be used as embeddings—a new vectorized representation for the network states and the ranked significance of nodes. These embeddings can be used in various research tasks in different areas, such as the vaccination problem, node-blocking, cloud computing, etc.

#### V. RELATED WORK

Misinformation has been studied vastly from historical [47], [48], political [49], sociological [18], [50], medical [51], [52], psychological [53] and computer science perspectives. Although the latter will also be our perspective, the other aspects possibly have profound effects on how computer science research on the matter unfolds as well, as shown in [54]. However, the link between other perspectives with computer science is yet to be thoroughly investigated.

From a computer science perspective, the research is mainly focused on how the problem and the concept of misinformation are defined [55]; analyzing how online misinformation spreads [18], [56], detecting and stopping its propagation [52], [57], [58]. In this work, we propose a model for misinformation propagation in social networks conforming to a game-theoretic model. Information diffusion on graphs has been studied using game-theoretic models previously. In [59], a framework based on evolutionary game-theoretic models on graphs has been proposed and tested on various synthetic and real networks. Yang *et al.* [60] propose and analyze an information spread model based on the diffusion of competitive information on graphs. The diffusion of rumor and misinformation based on game-theoretic models has been studied recently as well. Kumar and Geethakumari [61] create a model for misinformation spread. Their approach is different from ours in that they approach cooperation as a means to spread misinformation, which is the opposite of our approach. Li *et al.* [62] describe an evolutionary game with a punishment mechanism and a probabilistic

function to update node strategies. Xiao *et al.* [63] introduce internal and external factors and use them in a model of rumor propagation under a rumor/anti-rumor setting. Askarizadeh *et al.* [64], [65] explore an evolutionary model incorporating factors that affect rumor propagation and its control.

In addition, most of the systems that argue to successfully identify and stop misinformation from propagating record a time delay. This may indicate, from the attackers' perspective, that the damage would already be enough to influence as many people [66]. Hence, a preventive rather than a detection-based method may be more effective in defending against the problem. It is reported that large social network companies employ fake news and troll account detection techniques [67]; however, it is arguable that these platforms remain to be the top sources of misinformation as there is continuous research on these platforms [55]. In addition, due to their commercial and centralized setups, such platforms are unlikely to deliver transparency and objectiveness by being able to show interactions as is without incorporating algorithms that affect how the interactions are stored, distributed, and shown to people.

Manually identifying "fake news" has also been a motive for some organizations to create awareness. However, as these organizations curate content with manual labor most of the time, it is hard for them to keep track of emerging news stories and reach the level of breadth to cover the entire news ecosystem. In addition, their maintenance and development are done solely by a group of certain people, i.e., they are centralized, and most of them cannot guarantee transparency and prevent manipulation, for instance, in the form of cherry-picking in favor of some party.

## VI. CONCLUSION

In this work, we tackled the problem of misinformation propagation in OSNs from a game perspective. First, we approached the problem from the node-level point of view, where nodes were the actual players. We illustrated that the misinformation game constructed as a cooperative game on graphs displays the same characteristics that were explored in the literature. On the other hand, it has practical drawbacks, such as determining real values for variables such as benefit and cost described within the game dynamics. On the other hand, a more practical approach is possible with network-level players. We showed that a deep RL algorithm based on MADDPG can select an influential set of nodes in terms of misinformation propagation, and it gives promising results against various well-known algorithms such as CELF, pagerank, and node-centrality.

In future work, the explainability of the selections by the neural nets of RL agents could be studied to understand and implement better defense scenarios to stop misinformation dissemination. In addition, the behavior of the RL agents could be investigated in different types of networks and different tasks, domains, and different spread characteristics that are associated with the applications of node importance such as epidemiology, vaccination, cloud computing, and the Internet of Things.

## REFERENCES

- [1] N. Newman, R. Fletcher, A. Kalogeropoulos, D. Levy, and R. K. Nielsen, "Reuters Institute Digital News Report 2017," Reuters Inst., Oxford, U.K., Tech. Rep., 2017.
- [2] S. Flaxman, S. Goel, and J. M. Rao, "Filter bubbles, echo chambers, and online news consumption," *Public Opinion Quart.*, vol. 80, no. S1, pp. 298–320, 2016.
- [3] H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," *J. Econ. Perspect.*, vol. 31, no. 2, pp. 211–236, 2017.
- [4] C. F. Greer and D. A. Ferguson, "Using Twitter for promotion and branding: A content analysis of local television Twitter sites," *J. Broadcast. Electron. Media*, vol. 55, no. 2, pp. 198–214, May 2011.
- [5] G. Pressgrove, B. W. McKeever, and S. M. Jang, "What is contagious? Exploring why content goes viral on Twitter: A case study of the ALS ice bucket challenge," *Int. J. Nonprofit Voluntary Sector Marketing*, vol. 23, no. 1, Feb. 2018, Art. no. e1586.
- [6] J. Yin, S. Karimi, A. Lampert, M. Cameron, B. Robinson, and R. Power, "Using social media to enhance emergency situation awareness," in *Proc. 24th Int. Joint Conf. Artif. Intell.*, 2015, pp. 4234–4238.
- [7] K. Starbird and L. Palen, "'Voluntweeters': Self-organizing by digital volunteers in times of crisis," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, May 2011, pp. 1071–1080.
- [8] A. L. Schmidt *et al.*, "Anatomy of news consumption on Facebook," *Proc. Nat. Acad. Sci. USA*, vol. 114, no. 12, pp. 3035–3039, Mar. 2017.
- [9] J. L. Nelson and H. Taneja, "The small, disloyal fake news audience: The role of audience availability in fake news consumption," *New Media Soc.*, vol. 20, no. 10, pp. 3720–3737, Oct. 2018.
- [10] G. Szabó and G. Fáth, "Evolutionary games on graphs," *Phys. Rep.*, vol. 446, nos. 4–6, pp. 97–216, Jul. 2007.
- [11] E. Lieberman, C. Hauert, and M. A. Nowak, "Evolutionary dynamics on graphs," *Nature*, vol. 433, no. 7023, pp. 312–316, Jan. 2005.
- [12] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak, "A simple rule for the evolution of cooperation on graphs and social networks," *Nature*, vol. 441, pp. 502–505, May 2006.
- [13] M. Doebeli and C. Hauert, "Models of cooperation based on the Prisoner's dilemma and the snowdrift game," *Ecol. Lett.*, vol. 8, no. 7, pp. 748–766, Jun. 2005.
- [14] M. A. Nowak, "Five rules for the evolution of cooperation," *Science*, vol. 314, no. 5805, pp. 1560–1563, Dec. 2006.
- [15] V. P. Crawford, "Lying for strategic advantage: Rational and boundedly rational misrepresentation of intentions," *Amer. Econ. Rev.*, vol. 93, no. 1, pp. 133–149, 2003.
- [16] M. Zimdars and K. McLeod, *Fake News: Understanding Media and Misinformation in the Digital Age*. Cambridge, MA, USA: MIT Press, 2020.
- [17] V. Almeida, F. Filgueiras, and F. Gaetani, "Digital governance and the tragedy of the commons," *IEEE Internet Comput.*, vol. 24, no. 4, pp. 41–46, Jul. 2020.
- [18] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," *Science*, vol. 359, pp. 1146–1151, May 2018.
- [19] M. Glenski, T. Weninger, and S. Volkova, "Propagation from deceptive news sources who shares, how much, how evenly, and how quickly?" *IEEE Trans. Computat. Social Syst.*, vol. 5, no. 4, pp. 1071–1082, Dec. 2018.
- [20] J. M. Smith and G. R. Price, "The logic of animal conflict," *Nature*, vol. 246, no. 5427, pp. 15–18, Nov. 1973.
- [21] H. Ohtsuki and M. A. Nowak, "Direct reciprocity on graphs," *J. Theor. Biol.*, vol. 247, no. 3, pp. 462–470, Aug. 2007.
- [22] J. Leskovec and J. Mcauley, "Learning to discover social circles in ego networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 25, 2012, pp. 548–556.
- [23] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *J. Stat. Mech., Theory Exp.*, vol. 2008, no. 10, Oct. 2008, Art. no. P10008.
- [24] U. N. Raghavan, R. Albert, and S. Kumara, "Near linear time algorithm to detect community structures in large-scale networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 76, no. 3, Sep. 2007, Art. no. 036106.
- [25] E. D. Raj *et al.*, "Information granulation-based community detection for social networks," *IEEE Trans. Computat. Social Syst.*, vol. 8, no. 1, pp. 122–133, Feb. 2021.
- [26] D. Madeo and C. Mocenni, "Game interactions and dynamics on networked populations," *IEEE Trans. Autom. Control*, vol. 60, no. 7, pp. 1801–1810, Jul. 2015.
- [27] G. Como, F. Fagnani, and L. Zino, "Imitation dynamics in population games on community networks," *IEEE Trans. Control Netw. Syst.*, vol. 8, no. 1, pp. 65–76, Mar. 2021.

- [28] S. Rao, A. K. Verma, and T. Bhatia, "A review on social spam detection: Challenges, open issues, and future directions," *Expert Syst. Appl.*, vol. 186, Dec. 2021, Art. no. 115742.
- [29] P. K. Roy and S. Chahar, "Fake profile detection on social networking websites: A comprehensive review," *IEEE Trans. Artif. Intell.*, vol. 1, no. 3, pp. 271–285, Dec. 2020.
- [30] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *Proc. 9th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2003, pp. 137–146.
- [31] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance, "Cost-effective outbreak detection in networks," in *Proc. 13th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2007, pp. 420–429.
- [32] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, pp. 529–533, Feb. 2015.
- [33] R. S. Sutton, D. McAllester, S. Singh, and Y. Mansour, "Policy gradient methods for reinforcement learning with function approximation," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 12, 1999, pp. 1057–1063.
- [34] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller, "Deterministic policy gradient algorithms," in *Proc. Int. Conf. Mach. Learn.*, 2014, pp. 387–395.
- [35] T. P. Lillicrap *et al.*, "Continuous control with deep reinforcement learning," 2015, *arXiv:1509.02971*.
- [36] A. S. Bedi, S. Chakraborty, A. Parayil, B. M. Sadler, P. Tokekar, and A. Koppel, "On the hidden biases of policy mirror ascent in continuous action spaces," in *Proc. Int. Conf. Mach. Learn.*, 2022, pp. 1716–1731.
- [37] S. M. Kakade, *On the Sample Complexity of Reinforcement Learning*. London, U.K.: Univ. London, Univ. College London, 2003.
- [38] H. Kumar, A. Koppel, and A. Ribeiro, "On the sample complexity of actor-critic method for reinforcement learning with function approximation," 2019, *arXiv:1910.08412*.
- [39] G. Matheron, N. Perrin, and O. Sigaud, "The problem with DDPG: Understanding failures in deterministic environments with sparse rewards," 2019, *arXiv:1911.11679*.
- [40] R. Lowe, Y. I. Wu, A. Tamar, J. Harb, O. P. Abbeel, and I. Mordatch, "Multi-agent actor-critic for mixed cooperative-competitive environments," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 6379–6390.
- [41] J. Leskovec and J. McAuley, "Learning to discover social circles in ego networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 25, 2012, pp. 539–547.
- [42] A. L. Traud, P. J. Mucha, and M. A. Porter, "Social structure of Facebook networks," *Phys. A, Stat. Mech. Appl.*, vol. 391, no. 16, pp. 4165–4180, Aug. 2012.
- [43] M. Richardson, R. Agrawal, and P. Domingos, "Trust management for the semantic web," in *Proc. Int. Semantic Web Conf.* Cham, Switzerland: Springer, 2003, pp. 351–368.
- [44] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 855–864.
- [45] B. Perozzi, R. Al-Rfou, and S. Skiena, "DeepWalk: Online learning of social representations," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2014, pp. 701–710.
- [46] S. Ivanov and E. Burnaev, "Anonymous walk embeddings," in *Proc. Int. Conf. Mach. Learn.*, 2018, pp. 2186–2195.
- [47] J. M. Burkhardt, "History of fake news," *Library Technol. Rep.*, vol. 53, no. 8, pp. 5–9, 2017.
- [48] A. Amarasingam, *The Stewart/Colbert Effect: Essays Real Impacts Fake News*. McFarland, WI, USA: McFarland, 2014.
- [49] J. Tucker *et al.*, "Social media, political polarization, and political disinformation: A review of the scientific literature," *SSRN Electron. J.*, 2018, doi: [10.2139/ssrn.3144139](https://doi.org/10.2139/ssrn.3144139).
- [50] D. M. J. Lazer *et al.*, "The science of fake news," *Science*, vol. 359, no. 6380, pp. 1094–1096, 2018.
- [51] A. Kata, "A postmodern Pandora's box: Anti-vaccination misinformation on the Internet," *Vaccine*, vol. 28, no. 7, pp. 1709–1716, Feb. 2010.
- [52] L. Li *et al.*, "Characterizing the propagation of situational information in social media during COVID-19 epidemic: A case study on Weibo," *IEEE Trans. Computat. Social Syst.*, vol. 7, no. 2, pp. 556–562, Mar. 2020.
- [53] E. F. Loftus, "Planting misinformation in the human mind: A 30-year investigation of the malleability of memory," *Learn. Memory*, vol. 12, no. 4, pp. 361–366, Jul. 2005.
- [54] K. P. K. Kumar and G. Geethakumari, "Detecting misinformation in online social networks using cognitive psychology," *Hum.-Centric Comput. Inf. Sci.*, vol. 4, no. 1, pp. 1–22, Dec. 2014.
- [55] S. Kumar and N. Shah, "False information on web and social media: A survey," 2018, *arXiv:1804.08559*.
- [56] M. Del Vicario *et al.*, "The spreading of misinformation online," *Proc. Nat. Acad. Sci. USA*, vol. 113, no. 3, pp. 554–559, 2016.
- [57] V. Qazvinian, E. Rosengren, D. Radev, and Q. Mei, "Rumor has it: Identifying misinformation in microblogs," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2011, pp. 1589–1599.
- [58] G. Shrivastava *et al.*, "Defensive modeling of fake news through online social networks," *IEEE Trans. Computat. Social Syst.*, vol. 7, no. 5, pp. 1159–1167, Oct. 2020.
- [59] C. Jiang, Y. Chen, and K. R. Liu, "Graphical evolutionary game for information diffusion over social networks," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 4, pp. 524–536, Aug. 2014.
- [60] X. Yang, Z. Zhu, H. Yu, Y. Zhao, and L. Guo, "Evolutionary game dynamics of the competitive information propagation on social networks," *Complexity*, vol. 2019, pp. 1–11, Dec. 2019.
- [61] K. K. Kumar and G. Geethakumari, "Information diffusion model for spread of misinformation in online social networks," in *Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI)*, Aug. 2013, pp. 1172–1177.
- [62] D. Li, J. Ma, Z. Tian, and H. Zhu, "An evolutionary game for the diffusion of rumor in complex networks," *Phys. A, Stat. Mech. Appl.*, vol. 433, pp. 51–58, Sep. 2015.
- [63] Y. Xiao, D. Chen, S. Wei, Q. Li, H. Wang, and M. Xu, "Rumor propagation dynamic model based on evolutionary game and anti-rumor," *Nonlinear Dyn.*, vol. 95, no. 1, pp. 523–539, 2019.
- [64] M. Askarizadeh, B. T. Ladani, and M. H. Manshaei, "An evolutionary game model for analysis of rumor propagation and control in social networks," *Phys. A, Stat. Mech. Appl.*, vol. 523, pp. 21–39, Jun. 2019.
- [65] M. Askarizadeh and B. T. Ladani, "Soft rumor control in social networks: Modeling and analysis," *Eng. Appl. Artif. Intell.*, vol. 100, Apr. 2021, Art. no. 104198.
- [66] E. Thorson, "Belief echoes: The persistent effects of corrected misinformation," *Political Commun.*, vol. 33, no. 3, pp. 460–480, Jul. 2016.
- [67] N. Wingfield, M. Isaac, and K. Benner, *Google and Facebook Take Aim at Fake News Sites*, vol. 11. New York, NY, USA: New York Times, 2016, p. 12.



**Tolga Yilmaz** received the B.Sc. and M.Sc. degrees from Bilkent University, Ankara, Turkey, in 2013 and 2016, respectively, where he is currently pursuing the Ph.D. degree with the Department of Computer Engineering.

He is a Senior Researcher at the TÜBİTAK BİLGEM Cyber Security Institute, Ankara. His research interests include information retrieval, social networks, data mining, machine learning, and information security.



**Özgür Ulusoy** (Member, IEEE) received the Ph.D. degree in computer science from the University of Illinois at Urbana-Champaign, Champaign, IL, USA, in 1992.

He is currently a Professor at the Department of Computer Engineering, Bilkent University, Ankara, Turkey. He has published over 150 articles in archived journals and conference proceedings. His current research interests include web databases and web information retrieval, multimedia database systems, social networks, and cloud computing.

Dr. Ulusoy is a member of ACM.