Abstract—Social networks have become ubiquitous in our daily life, as such it has attracted great research interests recently. A key challenge is that it is of extremely large-scale with tremendous information flow, creating the phenomenon of "Big Data". Under such a circumstance, understanding information diffusion over social networks has become an important research issue. Most of the existing works on information diffusion analysis are based on either network structure modeling or empirical approach with dataset mining. However, the information diffusion is also heavily influenced by network users' decisions, actions and their socio-economic connections, which is generally ignored in existing works. In this paper, we propose an evolutionary game theoretic framework to model the dynamic information diffusion process in social networks. Specifically, we analyze the framework in uniform degree and non-uniform degree networks and derive the closed-form expressions of the evolutionary stable network states. Moreover, the information diffusion over two special networks, Erdős-Rényi random network and the Barabási-Albert scale-free network, are also highlighted. To verify our theoretical analysis, we conduct experiments by using both synthetic networks and real-world Facebook network, as well as real-world information spreading dataset of Twitter and Memetracker. Experiments shows that the proposed game theoretic framework is effective and practical in modeling the social network users' information forwarding behaviors.

Index Terms—Social networks, information diffusion, information spreading, game theory, evolutionary game,
Recently, as social networks, e.g., Facebook and Twitter, become more and more popular, some empirical analysis were conducted using large-scale datasets, including predicting the speed and range of information diffusion on Twitter [15], modeling the global influence of a node on the rate of diffusion on Twitter [16], illustrating the statistical mechanics of rumor spreading on Facebook [17]. Moreover, the information diffusion on overlaying social-physical networks was analyzed in [18]. The information diffusion analysis presented in this paper falls into the second category, i.e., our focus is on the analysis of dynamic diffusion process.

Most of existing works on diffusion analysis mainly rely on the social network structure and/or the collected data, while totally ignore the actions and decision making of users. However, the influence of users’ decisions, actions and socio-economic connections on information forwarding also plays an important role in the diffusion process. What does that actually mean? In essence, for information to diffuse, it relies on other users to forward the information they receive. However, it is not unconditional. One has to make a decision on whether or not to do so based on many factors, such as if the information is exciting or if his/her friends are interested in it, etc. This kind of interaction can often be modeled by using game theory [19]. Therefore, in this paper, we propose a game theoretic framework to analyze the information diffusion over social network. Specifically, we find that in essence the information diffusion process on social networks follows similarly the evolution process in natural ecological systems [20]. It is a process that evolves from one state at a particular instance to another when information is being forwarded and diffused around. Thus, we consider the graphical evolutionary game to model and analyze the social network users’ information forwarding strategies and the dynamic diffusion process.

The contributions of this paper can be summarized as follows.

1) We propose a graphical evolutionary game theoretic framework to model and understand information diffusion over social networks. The framework reveals the dynamics of information diffusion among users through analyzing their learning, interactions and decision making. Moreover, the framework not only can illustrate the dynamic process of the information diffusion, but also predict the final diffusion state.

2) Based on the graphical evolutionary game theoretic formulation, we analyze the dynamic diffusion process over uniform degree and non-uniform degree networks. The closed-form expressions for the evolutionary stable network states are obtained. Moreover, the information diffusion over two special networks, Erdős-Rényi random network and the Barabási-Albert scale-free network, are also studied.

3) We conduct experiments to validate our information diffusion analysis using both synthetic networks including uniform-degree network, Erdős-Rényi random network and the Barabási-Albert scale-free network, and real-world Facebook network. Moreover, we also use the real-world information spreading dataset from Memetracker and Twitter [21] to further verify the diffusion analysis.

The rest of this paper is organized as follows. We first give a brief overview of graphical evolutionary game formulation of information diffusion in Section II. Then, we analyze the dynamic information diffusion process over uniform and non-uniform degree networks in Section III and IV, respectively. Experiments results are shown in Section V and conclusions are drawn in Section VI.

II. GRAPHICAL EVOLUTIONARY GAME FORMULATION

In this section, we discuss the formulation of the information diffusion problem using graphical evolutionary game. We first introduce the basic concepts of graphical evolutionary game, and then elaborate how to formulate the information diffusion process using graphical evolutionary game theory.

A. Basic Concepts of Graphical Evolutionary Game

The traditional interpretation of game theory is that a game with some specific rule is played among a group of static players and a static Nash equilibrium (NE) can be achieved by analyzing the players’ payoff, utility function and the game rule. Evolutionary game theory (EGT) originated from ecological biology [20], instead, imagines that a game is played over and over again by biologically or socially conditioned players who are randomly drawn from large populations [22]. It studies the population shift and evolving process due...
to the influence of mutants. Meanwhile, EGT emphasizes more on the dynamics and stability of the whole population’s strategies, instead of only focusing on the property of the static NE. Recently, EGT has been widely used to model users’ behaviors in communication and networking area [23][24], including congestion control [25], cooperative sensing [26], cooperative peer-to-peer (P2P) streaming [27] and dynamic spectrum access [28], and also image processing area [29]. In these literatures, evolutionary game has been shown to be an effective approach to model the dynamic social interactions among users in a network.

EGT studies how a group of players converges to a stable equilibrium after a period of strategic interactions. Such a final equilibrium state is defined as the Evolutionarily Stable State (ESS). Let us consider an evolutionary game with $m$ strategies $X = \{1, 2, ..., m\}$ and payoff matrix, $U$, which is an $m \times m$ matrix, whose entry, $u_{ij}$, denotes the payoff for strategy $i$ versus strategy $j$. The system state of this game can be defined as $p = [p_1, ..., p_i, ..., p_m]^T$, where $p_i$ is the fraction of population using strategy $i$ and $\sum_{i=1}^{m} p_i = 1$. In such a case, the average mean payoff within a (sub-)population in state $q = [q_1, q_2, ..., q_m]^T$ against a population in state $p$ is $q^T U p$. A state $p^*$ is an ESS, if and only if for all different states $q \neq p$, $p^*$ satisfies follows [30]:

1) $q^T U^* \leq p^T U^* \leq p^* U q > q^T U q$, \hspace{1cm} (1)
2) if $q^T U^* = p^T U^*$, \hspace{1cm} (2)

The first condition guarantees that the average mean payoff of the population in a different state $q$ against the ESS $p^*$ does not exceed the average payoff of the population in $p^*$, which is equivalent to the Nash equilibrium condition. The second condition shows that, in case of equality in the equilibrium condition, the average payoff of state $q$ against itself is lower than that of state $p^*$ against $q$, which further guarantees the stability of the ESS $p^*$. From such a definition, we can see that even if a small fraction of players may not be rational and take out-of-equilibrium strategies, ESS is still a locally stable state. How to find the ESSs is an important issue in EGT. One common approach is to find the stable points of the system state dynamic $\dot{p} = dp/dt$ [22], i.e.,

$$p^* = \arg \left( \dot{p} = 0 \right).$$ \hspace{1cm} (3)

The classical evolutionary game theory considers a population in a complete graph. However, in many scenarios, players’ spatial locations may lead to an incomplete graph structure. Graphical evolutionary game theory is introduced to study the strategies evolution in such a finite structured population [31]. In graphical EGT, the “fitness” of a player is locally determined from interactions with all adjacent players, which is defined as [32]

$$\pi = (1 - \alpha) \cdot B + \alpha \cdot U,$$ \hspace{1cm} (4)

where $B$ is the baseline fitness, which represents the player’s inherent property. For example, in a social network, a user’s
baseline fitness can be interpreted as his/her own interests on the released news. \( U \) is the player’s payoff which is determined by the predefined payoff matrix and the graph structure. The parameter \( \alpha \) represents the selection intensity, i.e., the relative contribution of the game to fitness. The case \( \alpha \to 0 \) represents the limit of weak selection \[33\], while \( \alpha = 1 \) denotes strong selection, where fitness equals payoff. The ESS of graphical EGT is also of importance, which is usually analyzed under different strategy updating rules, including birth-death (BD), death-birth (DB) and imitation (IM) introduced as follows \[34\].

- BD update rule: a player is chosen for reproduction with the probability being proportional to fitness (Birth process). Then, the chosen player’s strategy replaces one neighbor’s strategy with uniform probability (Death process), as shown in Fig. 2(a).
- DB update rule: a random player is chosen to abandon his/her current strategy (Death process). Then, the chosen player adopts one of his/her neighbors’ strategies with the probability being proportional to their fitness (Birth process), as shown in Fig. 2(b).
- IM update rule: each player either adopts the strategy of one neighbor or remains with his/her current strategy, with the probability being proportional to fitness, as shown in Fig. 2(c).

These strategy update rules are also from the evolutionary biology filed, which are used to model the resident/mutant evolution process. Through analyzing the system dynamics under some specific strategy update rule, the final ESSs can be found at the stable points of the dynamics. In summary, graph structure, players, strategy, fitness (payoff) and ESS are five basic elements of a graphical evolutionary game. In the following, we will illustrate how to formulate the dynamic information diffusion process in a social network using graphical EGT.

### B. Graphical Evolutionary Game Formulation

A social network is usually illustrated by a graph, where each node represents a user and the edge represents the relationship between users. The users can be either human in a social network or websites on the internet, while the relationship can be either friendship between users or hyperlink between webpages. When some new information is originated from one user, the information may be propagated over the network depending on other users’ actions: to forward the information or not. For each user, whether to forward the information is determined by several factors, including the user’s own interest on this information and his/her neighbor’s actions in the sense that if all his/her neighbors forward the information, the user may also forward the information with relatively high probability. In such a case, the users’ actions are coupled with each other through their social interactions. This is very similar to the player’s strategy update in the aforementioned graphical EGT, where players’ strategies are also influenced with each other through the graph structure.

Compared with the strategy updating in the graphical EGT, the information diffusion over social network shares fundamental similarities. A user in a social network can be regarded as a player in a graphical evolutionary game. And each user’s two possible actions, i.e., to forward or not forward, are corresponding with two strategies

\[
\begin{align*}
S_f, & \quad \text{forward the information}, \\
S_n, & \quad \text{not forward the information}.
\end{align*}
\]

Meanwhile, users’ payoff matrix can be written as

\[
\begin{bmatrix}
S_f & S_n \\
\frac{u_{ff}}{u_{fn}} & \frac{u_{fn}}{u_{nn}}
\end{bmatrix}
\]  

(6)

where a symmetric payoff structure is considered, i.e., when a user with strategy \( S_f \) meets a user with strategy \( S_n \), each of them receives the same payoff \( u_{fn} \). Moreover, we assume that the payoff has been normalized within interval \((0, 1)\), i.e., \(0 < u_{ff}, u_{fn}, u_{nn} < 1\). The physical meaning of the payoff can be either the popularity of a user in a social network or the hit rate of a website. Note that under different application scenarios, the values of the payoff matrix may be different.

For example, if the information is related to recent hot topics and forwarding of the information can attract more attentions from other users or website, the payoff matrix should have the following characteristic: \( u_{ff} \geq u_{fn} \geq u_{nn} \). On the other hand, if the information is about useless advertisements, the payoff matrix would exhibit \( u_{fn} \geq u_{ff} \). Furthermore, if the information is supposed to be shared only within a circle, i.e., a small group with same interest, the payoff matrix could exhibit \( u_{fn} \geq u_{nn} \). It is also worth mentioning that the payoff matrix defined in (6) is similar to that of the coordination game, where two users making the same actions can obtain symmetric higher payoff than making different actions \[35\].

Based on the above definitions of strategy and payoff function, we can further define the fitness of users with strategy \( S_f \) and strategy \( S_n \), respectively. According to \[4\], for a user with strategy \( S_f \) and \( k \) neighbors, among which \( k_f \) users adopting the same strategy \( S_f \) and \( k - k_f \) users adopting strategy \( S_n \), his/her fitness can be written as

\[
\pi_f(k, k_f) = (1 - \alpha) + \alpha [k_f u_{ff} + (k - k_f) u_{fn}],
\]

(7)
where the baseline fitness is normalized as 1. Similarly, for a user with strategy $S_n$ and $k$ neighbors, among which $k_n$ users adopting the same strategy $S_n$ and $k-k_n$ users adopting strategy $S_f$, his/her fitness can be written as

$$
\pi_n(k, k_n) = (1 - \alpha) + \alpha [k_n u_{nn} + (k - k_n)u_{fn}].
$$

Eqn. (8) implies that the state of the whole social network can be described by only two variables, $p_f$ and $p_{ff}$. In the following, we will first analyze the dynamics of $p_f$ and $p_{ff}$. Based on the dynamics analysis, the evolutionary stable network state can be achieved.

### A. Evolution Dynamics of Network State

In order to derive the network ESS, we need to analyze the evolution dynamics of the network state. As discussed above, the network state can be represented by parameters $p_f$ and $p_{ff}$, i.e., the network user state and network edge state.

In such a case, we will first derive the evolution dynamics of $p_f$ and $p_{ff}$ under the IM update rule, to understand the information diffusion process over the social network. Then, through analyzing the stable point of the network dynamics $\dot{p}_f$ and $\dot{p}_{ff}$, the final network ESS can be achieved.

According to the IM update rule, a user adopting strategy $S_f$, i.e., forward the information, is selected for imitation with probability $p_f$. Fig. 3 illustrates a scenario where the center user adopts strategy $S_f$, and among his/her $k$ neighbors, there are $k_f$ users adopting the same strategy $S_f$ and $k-k_f$ users adopting the opposite strategy $S_n$. The probability of such a configuration is

$$
\theta_f(k,k_f) = \binom{k}{k_f} p_{ff}^{k_f} (1-p_{ff})^{k-k_f}.
$$

Eqn. (12) implies that the state of the whole social network can be described by only two variables, $p_f$ and $p_{ff}$. In the following, we will first analyze the dynamics of $p_f$ and $p_{ff}$. Based on the dynamics analysis, the evolutionary stable network state can be achieved.

### III. Information Diffusion over Uniform Degree Network

In this section, we study the information diffusion process over networks with uniform degree using graphical EGT. In the uniform scenario, an $N$-user social network based on a homogenous graph with general degree $k$ is considered. When there is new information originated from some user, other users updates their forward strategies, i.e. either forward the information or not (see $S_f$ or $S_n$). As for the users’ strategy update rule, we model it using the aforementioned IM update rule. Note that all the analysis in the section can be easily extended to the BD and DB update rules. Let us define the network user state as $(p_f, p_n)$, which denotes the percentages of users choosing strategies $S_f$ and $S_n$ among the whole population, respectively. Our ultimate goal is to derive the evolutionary stable network user state $(p^*_f, p^*_n)$, which represents the final diffusion status of the information.

In order to illustrate the network state evolving process on a graph, the edge information of the graph is also required. Let us further define the network edge state as $(p_{ff}, p_{fn}, p_{nn})$, where $p_{ff}$ and $p_{nn}$ represent the percentages of edges on which both users choose same strategy $S_f$ and $S_n$, respectively; $p_{fn}$ represents the percentage of edges on which one user chooses to forward the information while the other chooses not. Moreover, let $p_{fn}$ denote the percentage of a user’s neighbors using strategy $S_f$, given the user is using strategy $S_n$. Similarly, we have $p_{ff}$, $p_{nn}$ and $p_{fn}$. In such a case, according to the basic probability theory, we can summarize the relationship of these defined variables as

$$
\pi_f = (1 - \alpha) + \alpha [k_f u_{ff} + (k - k_f)u_{fn}].
$$

Among the center user’s $k$ neighbors, for the neighbors adopting strategy $S_f$, each of them has $(k-1)p_{nn}$ neighbors using strategy $S_n$, and $(k-1)p_{ff}+1$ neighbors using strategy $S_f$ averagely, where “1” means the center user with strategy $S_f$, as shown by the dashed lines in the left part of Fig. 3. In
such a case, the average fitness of the center user’s neighbors who adopt strategy \( S_f \), i.e., forward the information, is

\[
\pi_{fj} = (1 - \alpha) + \alpha \left( (k-1)p_{n|f}u_{fn} + [(k-1)p_{f|f} + 1]u_{ff} \right).
\] (14)

Similarly, for the center user’s neighbors adopting strategy \( S_n \), each of them has \( (k-1)p_{n|n} \) neighbors using strategy \( S_n \) and \( (k-1)p_{f|n} + 1 \) neighbors using strategy \( S_f \). Thus, the average fitness of the center user’s neighbors who adopt strategy \( S_n \), i.e., not forward the information, is

\[
\pi_{nj} = (1 - \alpha) + \alpha \left( (k-1)p_{n|n}u_{nn} + [(k-1)p_{f|n} + 1]u_{fn} \right).
\] (15)

According to the IM update rule, the probability of a user updating his/her strategy is proportional to the fitness of the corresponding strategy. In such a case, as shown in the middle part of Fig. 4, the probability that the center user updates his/her current strategy \( S_f \) by strategy \( S_n \) is

\[
P_{f \rightarrow n}(k_f) = \frac{(k-k_f)\pi_{n|f}}{k_f\pi_{f|f} + (k-k_f)\pi_{n|f} + \pi_n}.
\] (16)

Therefore, as shown in right part of Fig. 4, the center user deviating from using strategy \( S_f \) leads to the percentage of users adopting strategy \( S_f \), \( p_f \), decreasing by \( 1/N \) with probability

\[
\text{Prob}(\Delta p_f = -\frac{1}{N}) = \frac{1}{N} \sum_{k_f=0}^{k} p_f \theta_f(k, k_f) P_{f \rightarrow n}(k_f)
\]

\[
= p_f \sum_{k_f=0}^{k} \left( \frac{k}{k_f} \right) p_{f|f}(1 - p_{f|f})^{k-k_f} \frac{(k-k_f)\pi_{n|f}}{k_f\pi_{f|f} + (k-k_f)\pi_{n|f} + \pi_n}.
\] (17)

Meanwhile, the center user deviating from using strategy \( S_f \) also leads to the edges on which both users adopting strategy \( S_f \) decreasing by \( k_f \), which means that \( p_{ff} \) decreases by \( \frac{2k_f}{k} \) with probability

\[
\text{Prob}(\Delta p_{ff} = -\frac{2k_f}{kN}) = p_f \left( \frac{k}{k_f} \right) p_{f|f}(1 - p_{f|f})^{k-k_f} \frac{(k-k_f)\pi_{n|f}}{k_f\pi_{f|f} + (k-k_f)\pi_{n|f} + \pi_n}.
\] (18)

Similar analysis can be applied to the user adopting strategy \( S_n \). According to the IM update rule, a user adopting strategy \( S_n \) is selected for imitation with probability \( p_n = 1 - p_f \). As shown in Fig. 4, among its \( k \) neighbors, suppose that there are \( k_n \) users adopting the same strategy \( S_n \) and \( k-k_n \) users adopting the opposite strategy \( S_f \). The probability of such a configuration is

\[
\theta_n(k, k_n) = \left( \frac{k}{k_n} \right) p_{n|n}(1 - p_{n|n})^{k-k_n}.
\] (19)

Under such a configuration, the fitness of the center user who adopts strategy \( S_n \) is

\[
\pi_n = (1 - \alpha) + \alpha \left( k_n u_{nn} + (k-k_n) u_{fn} \right).
\] (20)

Among his/her \( k \) neighbors, the average fitness of neighbors who adopt strategy \( S_f \) is

\[
\pi_{f|n} = (1 - \alpha) + \alpha \left( (k-1)p_{f|f}u_{ff} + [(k-1)p_{f|n} + 1]u_{fn} \right),
\] (21)

and the fitness of node adopting strategy \( S_n \) is

\[
\pi_{n|n} = (1 - \alpha) + \alpha \left( (k-1)p_{f|n}u_{fn} + [(k-1)p_{n|f} + 1]u_{nn} \right).
\] (22)

In such a case, as shown in the middle part of Fig. 4, the probability that the center user updates his/her strategy \( S_f \) by strategy \( S_n \) is

\[
P_{n \rightarrow f}(k_n) = \frac{(k-k_n)\pi_{f|n}}{k_n\pi_{n|n} + (k-k_n)\pi_{f|n} + \pi_n}.
\] (23)

Therefore, as shown in right part of Fig. 4, the center user deviating from using strategy \( S_n \) leads to the percentage of users adopting strategy \( S_f \), \( p_f \), increasing by \( 1/N \) with probability

\[
\text{Prob}(\Delta p_f = \frac{1}{N}) = \sum_{k_n=0}^{k} (1 - p_f) \theta_n(k, k_n) P_{n \rightarrow f}(k_n)
\]

\[
= (1 - p_f) \sum_{k_n=0}^{k} \left( \frac{k}{k_n} \right) p_{n|n}(1 - p_{n|n})^{k-k_n} \frac{(k-k_n)\pi_{f|n}}{k_n\pi_{n|n} + (k-k_n)\pi_{f|n} + \pi_n}.
\] (24)

Meanwhile, the deviation also leads to the edges on which both users adopting strategy \( S_f \) increasing by \( k - k_n \), thus,
we have
\[
\text{Prob}\left(\Delta p_{ff} = \frac{k - k_n}{kN/2}\right) = (1 - p_f) \frac{(k - k_n) p_{nn}^k}{kN\pi_n} (1 - p_n)\left(k - k_n\right) + O(\alpha^2),
\] (25)
Combining (17) and (24), we have the probabilities of \(p_f\) increasing \(1/N\) and decreasing \(1/N\), respectively. In such a case, we can calculate the expected variation of \(p_f\) per unit time, which is defined as the dynamic of \(p_f\), as
\[
\dot{p}_f = \frac{1}{N} \text{Prob}\left(\Delta p_f = \frac{1}{N}\right) - \frac{1}{N} \text{Prob}\left(\Delta p_f = -\frac{1}{N}\right)
= \frac{\alpha k(k - 1)p_{fn}}{2N(k + 1)^2} \left(\gamma_1 u_{nn} + 2u_{fn} + \gamma_3 u_{ff}\right) + O(\alpha^2),
\] (26)
where the second equality is according to Taylor’s Theorem and weak selection assumption with \(\alpha\) goes to zero [36], and the parameters \(\gamma_1, \gamma_2, \text{and} \gamma_3\) are given as follows:
\[
\begin{align*}
\gamma_1 &= -p_{nn}[(k - 1)(p_{nn} + p_{ff}) + 3], \\
\gamma_2 &= p_{nn} - p_{ff} + (p_{ff} - p_{fn})[(k - 1)(p_{nn} + p_{ff}) + 2], \\
\gamma_3 &= p_{ff}[(k - 1)(p_{nn} + p_{ff}) + 3].
\end{align*}
\] (27)
Under the weak selection, the payoff obtained from the interactions is considered as limited contribution to the overall fitness of each player. The results derived from weak selection often remain as valid approximations for larger selection strength [33]. Moreover, the weak selection has a long tradition in theoretical biology. [37]. Moreover, the weak selection assumption can help achieve a close-form analysis of diffusion process and better reveal how the strategy diffuses over the network. From (26), we can see that the dynamics \(\dot{p}_f\) is an increasing function in terms of \(u_{ff}\), i.e., the payoff of both users forwarding the released information. The corresponding physical meaning is that when the higher payoff can be obtained by forwarding the information, the increasing rate of \(p_f\) will also be higher. On the other hand, if not forwarding the information can gain higher payoff, the increasing rate of \(p_f\) will be lower, which is just the reason why \(\dot{p}_f\) is a decreasing function in terms of \(u_{nn}\), i.e., the payoff of both users do not forward the information. Similarly, by combining (18) and (25), we have the probabilities of \(p_{ff}\) increasing \(2k_f/kN\) and decreasing \(2(k - k_n)/kN\), respectively. Thus, we can calculate the expected variation of \(p_{ff}\) per unit time, which is defined as the dynamic of \(p_{ff}\), as
\[
\dot{p}_{ff} = \sum_{k_f = 0}^{k} \frac{-2k_f}{kN} \text{Prob}\left(\Delta p_{ff} = -\frac{2k_f}{kN}\right)
+ \sum_{k_n = 0}^{k} \frac{2(k - k_n)}{kN} \text{Prob}\left(\Delta p_{ff} = \frac{2(k - k_n)}{kN}\right)
= \frac{p_{fn}}{(k + 1)N} \left(1 + (k - 1)(p_{fn} - p_{ff})\right) + O(\alpha).
\] (30)

B. ESS of Information Diffusion

As discussed in Section II-A, at the ESS, the network state dynamics achieve stable point, i.e., \(\dot{p}_f = 0\) and \(\dot{p}_{ff} = 0\). By setting \(\dot{p}_f = 0\), we have
\[
p_{ff} - p_{fn} = \frac{1}{k - 1}.
\] (31)
In such a case, according to (9)-(11) and (31), all the network state parameters can be expressed only by \(p_f\) as follow:
\[
\begin{align*}
p_{ff} &= p_f + \frac{1}{k - 1} (1 - p_f), \\
p_{fn} &= \frac{k - 2}{k - 1} p_f, \\
p_{ff} &= \frac{k - 2}{k - 1} (1 - p_f), \\
p_{fn}^* &= 1 - \frac{k - 2}{k - 1} p_f.
\end{align*}
\] (32-35)
Then, by substituting (32)-(35) and (27)-(29) into (26), we can obtain the dynamic of the network user state, \(\dot{p}_f\), as
\[
\dot{p}_f = \frac{\alpha k(k - 2)(k + 3)}{2N(k - 1)(k + 1)^2} p_f (1 - p_f) (ap_f - b),
\] (36)
where the parameters \(a\) and \(b\) are given as follows:
\[
\begin{align*}
a &= (k - 2)(u_{ff} - 2u_{fn} + u_{nn}), \\
b &= (k - 1)u_{nn} - (k - 2)u_{fn} - u_{ff}.
\end{align*}
\] (37-38)
At the ESS, we have \(\dot{p}_f = 0\). Therefore, there are three evolutionary stable network states:
\[
p_f^* = 0, 1 \text{ and } \frac{b}{a}.
\] (39)
The following theorem shows the conditions of these stable network states.

**Theorem 1:** In an \(N\)-user social network which can be characterized by a graph with uniform degree \(k\), if each user updates his/her information forward strategy using the IM update rule, the evolutionary stable network states can be summarized as follows:
\[
p_f^* = \begin{cases} 
1, & \text{if } u_{ff} > u_{fn} > u_{nn}, \\
0, & \text{if } u_{nn} > u_{fn} > u_{ff}, \\
\frac{(k - 2)u_{fn} + u_{ff} - (k - 1)u_{nn}}{2u_{fn} - (k - 2)(u_{ff} - u_{nn})}, & \text{else}.
\end{cases}
\] (40)

**Proof:** Considering (26) and (30) as a nonlinear dynamic system, we can examine whether the three stable points are evolutionary stable network states through analyzing the Jacobian matrix of the dynamics as follows:
\[
J = \begin{bmatrix}
\partial \dot{p}_f/\partial p_f & \partial \dot{p}_f/\partial p_{fn} & \partial \dot{p}_f/\partial p_{ff} \\
\partial \dot{p}_{ff}/\partial p_f & \partial \dot{p}_{ff}/\partial p_{fn} & \partial \dot{p}_{ff}/\partial p_{ff}
\end{bmatrix}.
\] (41)
The evolutionary stability requires that \(\det(J) > 0\) and \(\text{tr}(J) < 0\) [20]. By substituting (26) and (30) into (41), we can obtain the three conditions for \(p_f^* = 0, 1\) and \(\frac{b}{a}\), respectively. Due to page limitation, the detailed derivations are omitted here.

Among these three ESSs, \(p_f^* = 0\) and \(1\) are two extreme stable states, representing no user forwarding the information and all users forwarding the information, respectively. According
IV. INFORMATION DIFFUSION OVER NON-UNIFORM DEGREE NETWORK

In this section, we study the information diffusion process over networks with non-uniform degree. In the non-uniform scenario, we consider an $N$-user social network based on a graph whose degree exhibits distribution $\lambda(k)$. This distribution means that when randomly choosing one user on the network, the probability of the chosen user with $k$ neighbors is $\lambda(k)$. In such a case, the average degree of the network is

$$\bar{k} = \sum_{k=0}^{\infty} \lambda(k)k.$$  \hfill (43)

Note that we do not take degree correlation into account, i.e., the degrees of all users are independent of each other. Similar to the analysis in Section III, when some new information is released, all users update their information forwarding strategies in a spontaneous manner. In order to give more insights on different evolutionary update rules, in this section, we model the users’ information forwarding strategy update using the aforementioned BD update rule. In the following, we first analyze the general case of information diffusion over non-uniform degree networks. Then, we focus on two special cases, i.e., two typical random networks, Erdős-Rényi random network and Barabási-Albert scale-free networks.

A. General Case

In the analysis of non-uniform degree network, to ensure the derivation of ESS is mathematically realizable, we assume that the network user state $p_f$ and the network edge states $p_{ff}$, $p_{fn}$ and $p_{nn}$, do not depend on the degree. With such an assumption, eqns. (9) and (11) still hold for the non-uniform degree network, i.e., the network state can still be described by two variables $p_f$ and $p_{ff}$. As a matter of fact, it has been shown that the dependence of the network state on degree is weak [38]. Similar to the analysis in Section III, in order to find the evolutionary stable network state, we need to first derive the evolving dynamics of $p_f$ and $p_{ff}$.

According to the BD update rule, a user is chosen with the probability being proportional to fitness. Then, the chosen user’s strategy replaces one of his/her neighbors’ strategies uniformly. As shown in Fig.9, suppose the chosen user is adopting strategy $S_f$ and with degree $l$. Among his/her neighbors, there are $l_f$ users adopting the same strategy $S_f$ and $l - l_f$ users adopting the opposite strategy $S_n$. In such a case, the probability that the chosen user’s strategy $S_f$ replaces one of his/her neighbors adopting strategy $S_n$ is

$$P_{n \rightarrow f}(l, l_f) = p_f \theta_f(l, l_f) \frac{\pi_f(l, l_f)}{\pi} l - l_f,$$  \hfill (44)

where $\theta_f(l, l_f)$ is the probability of the configuration given in (12), $\pi_f(l, l_f)$ is the fitness of the chosen user given in (7) and $\pi$ is the expected fitness of all users which can be calculated
as follows:

\[
\pi = N p_f \sum_{k=0}^{+\infty} \lambda(k)(k p_{f|f} u_{ff} + k(1-p_{f|f}) u_{fn}) + N p_n \sum_{k=0}^{+\infty} \lambda(k)(k p_{n|n} u_{nn} + k(1-p_{n|n}) u_{fn})
\]

\[
= N(p_{f|f} u_{ff} + p_{f|n} u_{fn} + p_{n|n} u_{nn}). \quad (45)
\]

Therefore, the percentage of users adopting strategy \( S_f, p_f \), increases by \( 1/N \) with expected probability

\[
\text{Prob} \left( \Delta p_f = \frac{1}{N} \right) = \sum_{l=0}^{+\infty} \lambda(l) \sum_{l_f=0}^{l} p_f \theta_f(l, l_f) \frac{\pi_f(l, l_f)}{\pi} (l-l_f).
\]

(46)

Suppose the replaced user, who was adopting strategy \( S_n \), is with degree \( l' \). After strategy update, the edges on which both users adopting strategy \( S_f \) will increase by \( (l'-1)p_{f|f} + 1 \). In such a case, we have

\[
\text{Prob} \left( \Delta p_f = \frac{1}{N} \right) = \frac{1}{N} \sum_{l=0}^{+\infty} \lambda(l) \sum_{l_f=0}^{l} p_f \theta_f(l, l_f) \frac{\pi_f(l, l_f)}{\pi} (l-l_f).
\]

(47)

where \( k \) is the average degree of the network given in (43).

Similarly, suppose the chosen user is adopting strategy \( S_n \) and with degree \( m \) as shown in Fig. [4]. Among his/her neighbors, there are \( m_n \) users adopting the same strategy \( S_f \) and \( m-m_n \) users adopting the opposite strategy \( S_n \). In such a case, according to the BD update rule, the percentage of users adopting strategy \( S_f, p_f \), decreases by \( 1/N \) with expected probability

\[
\text{Prob} \left( \Delta p_f = \frac{1}{N} \right) = \sum_{m=0}^{+\infty} \lambda(m) \sum_{m_f=0}^{m} (1-p_f) \theta_n(m, m_f) \frac{\pi_n(m, m_f)}{\pi} m_f/m.
\]

(48)

where \( \theta_n(k, k_f) \) is given in (19), \( \pi_n(k, k_f) \) is given in (8) and \( \pi \) is given in (45). Suppose the replaced user, who was adopting strategy \( S_f \), is with degree \( m' \). After strategy update, the edges on which both users adopting strategy \( S_f \) will decrease by \( (m'-1)p_{f|f} \). In such a case, we have

\[
\text{Prob} \left( \Delta p_f = \frac{1}{N} \right) = \frac{1}{N} \sum_{m=0}^{+\infty} \lambda(m) \sum_{m_f=0}^{m} (1-p_f) \theta_n(m, m_f) \frac{\pi_n(m, m_f)}{\pi} m_f/m.
\]

(49)

Combining (46) and (48), we have the dynamics of \( p_f \) as

\[
\dot{p}_f = \frac{1}{N} \text{Prob} \left( \Delta p_f = \frac{1}{N} \right) - \frac{1}{N} \text{Prob} \left( \Delta p_f = -\frac{1}{N} \right) = \frac{\alpha(k-1)p_n}{\bar{k}N} \left[ p_{f|f}(u_{ff} - u_{fn}) - p_{n|n}(u_{nn} - u_{fn}) \right] + O(\bar{k}^2),
\]

(50)

where \( \bar{k} \) is the average degree of the network given in (43). Combining (47) and (49), we have the dynamics of \( p_{ff} \) as

\[
\dot{p}_{ff} = \frac{(\bar{T} - 1)p_{f|f} + 1}{N} \cdot \text{Prob} \left( \Delta p_f = \frac{1}{N} \right) - \frac{(m'-1)p_{f|f}}{N} \cdot \text{Prob} \left( \Delta p_f = -\frac{1}{N} \right), \quad (51)
\]

where \( \bar{T} \) and \( \bar{m}' \) are the average degrees of the replaced users during the strategy update, which are different with the average degree of the whole network, \( \bar{k} \). If a link is selected at random, the degree distribution of the node on the specific link is not \( \lambda(k) \) but rather \( \sum_{k=0}^{+\infty} k \lambda(k)\). Thus, the average degree of the replaced users is

\[
\bar{T} = \bar{m}' = \sum_{k=0}^{+\infty} k \lambda(k) = \frac{\bar{k}^2}{\bar{k}}, \quad (52)
\]

where \( \bar{k}^2 = \sum_{k=0}^{+\infty} k^2 \lambda(k) \) is the expectation of \( k^2 \). Based on the dynamics of \( p_f \) and \( p_{ff} \) in (50) and (51), we can obtain the following theorem.

Theorem 2: In an \( N \)-user social network which can be characterized by a graph with degree distribution \( \lambda(k) \), if each user updates his/her information forward strategy using the BD update rule, the evolutionary stable network states can be summarized as follows:

\[
p_f^* = \begin{cases} 
1, & \text{if } u_{ff} > u_{fn} > u_{nn}, \\
0, & \text{else}.
\end{cases}
\]

(53)

Proof: Similar to the proof of Theorem 1, the evolutionary stable network state \( p_f^* \) can be derived by solving \( \dot{p}_f = 0 \) and \( \dot{p}_{ff} = 0 \) in (51) and (52). Meanwhile, the conditions in (53) can be obtained by analyzing the Jacobian matrix of \( p_f \) and \( p_{ff} \). Due to page limitation, the detailed derivation is omitted.

B. Special Cases

In this subsection, we discuss two special cases of the non-uniform degree networks, Erdős-Rényi random network [40] and Barabási-Albert scale-free network [41]. For the Erdős-Rényi random (ER) network, the degree follows a Poisson distribution, i.e.,

\[
\lambda_{ER}(k) = \frac{e^{-\frac{\bar{k}k}{k!}}}{k!} \quad \text{and} \quad \bar{k}^2 = \bar{k} \bar{k} + 1. \quad (54)
\]

In such a case, according to Theorem 2, the ESS of Erdős-Rényi random network is

\[
p_{f|ER}^* = \begin{cases} 
1, & \text{if } u_{ff} > u_{fn} > u_{nn}, \\
0, & \text{else}.
\end{cases}
\]

(55)

For the Barabási-Albert scale-free (BA) network, the degree follows a power law distribution, i.e.,

\[
\lambda_{BA}(k) \propto k^{-\xi} \quad \text{and} \quad \bar{k}^2 = \bar{k} \log N^2 / 4 \quad \text{when } \xi = 3. \quad (56)
\]

In such a case, according to Theorem 2, the ESS of Barabási-Albert scale-free network is

\[
p_{f|BA}^* = \begin{cases} 
1, & \text{if } u_{ff} > u_{fn} > u_{nn}, \\
0, & \text{else}.
\end{cases}
\]

(57)
V. Experiments

In this section, we conduct experiments to verify the information diffusion analysis. First, we simulate the information diffusion process on synthetic networks. Then, we conduct information diffusion experiment on real-world networks, i.e., the Facebook social network. Finally, we use real-world information diffusion dataset to further verify the diffusion analysis.

A. Synthetic Networks

In the experiment of synthetic networks, we generate three kinds of networks to simulate the information diffusion process:

- the uniform-degree network;
- the Erdős-Rényi random network;
- the Barabási-Albert scale-free network.

For each network, we generate 1000 users and initialize each user with random strategy: \( S_f \) or \( S_n \). In the simulations, four kinds of payoff matrices are considered as follows:

- PM 1: \( u_{ff} > u_{fn} > u_{nn} \)
  \[
  S_f = \begin{pmatrix} 0.8 & 0.6 \\ 0.6 & 0.4 \end{pmatrix}
  \]
  (58)

- PM 2: \( u_{fn} > u_{ff} > u_{nn} \)
  \[
  S_f = \begin{pmatrix} 0.6 & 0.8 \\ 0.8 & 0.4 \end{pmatrix}
  \]
  (59)

- PM 3: \( u_{fn} > u_{nn} > u_{ff} \)
  \[
  S_f = \begin{pmatrix} 0.4 & 0.8 \\ 0.8 & 0.6 \end{pmatrix}
  \]
  (60)

- PM 4: \( u_{nn} > u_{fn} > u_{ff} \)
  \[
  S_f = \begin{pmatrix} 0.4 & 0.6 \\ 0.6 & 0.8 \end{pmatrix}
  \]
  (61)

Fig. 5 shows the experiment results for the uniform-degree network under different average degrees and payoff matrices. The theoretical results is calculated from (40) directly, while the simulation results are obtained by simulating the IM strategy update rule over the generated network. For each simulation run, the strategy update steps are repeated until the network reaches the stable network state. Meanwhile, the network structure is re-generated every 500 runs to prevent any spurious results based on one particular realization of a specific network type. From Fig. 5 we can see that all the simulation results are consistent with the theoretical results, which verifies the correctness of our diffusion analysis and the conclusion in Theorem 1. Moreover, different settings of the payoff matrix can lead to different evolutionary stable network states, while the network degree variations have little influence on the stable network states. Such a structure-free phenomenon is consistent with our approximation analysis for the large-degree case after Theorem 1, especially in (42).

Fig. 6 and 7 show the experiment results for the non-uniform degree networks under different average degrees and payoff matrices, including the Erdős-Rényi random network and the Barabási-Albert scale-free network. The theoretical results are calculated from (55) and (57) directly, while the simulation
results are obtained by simulating the BD strategy update rule over the two generated networks. We can see that the all the simulation results agree well with the theoretical results. In Fig. 7, the small gap for the Barabási-Albert scale-free networks is due to the fact that there is weak dependence between the network state and the network degree, while we neglected such dependence in the diffusion analysis. Nevertheless, we can see that the gap is relatively small and indeed negligible.

B. Real-World Networks

In the experiment of real-world network, we evaluate the information diffusion process over Facebook social network [42]. The Facebook dataset contains totally 4039 users and 88234 edges, where the edge means the connection between two users [4]. The abstract graph of the Facebook network is depicted in Fig. 8 where the nodes with large degrees, i.e., the users with large number of friends, are highlighted. Meanwhile, we also plot the degree characteristics of the Facebook network in Fig. 9 including the cumulative distribution (CDF) and probability density (PDF) of the network users' degrees. From the CDF, we can see that the users' degrees vary from 1 to 300 and 90 percents of the users are with less than 100 degrees. From the PDF where the axes are in logarithm scale, we can see that the Facebook network degree exhibits the scale-free phenomenon [41].

The Facebook dataset contains ten ego-networks, which means that there are ten subgraphs. In the experiment, we simulate the information diffusion process over the ten subgraphs respectively. Fig. 10 shows the experiment results under different payoff matrix settings. The theoretical results are calculated from (53), while the simulation results are obtained by simulating the BD strategy update rule over the ten subgraphs. It can be seen that the simulation results match well with the theoretical results for all ten subgraphs. The small gaps for some figures, e.g., the 10-th subgraph, are mainly due to the neglected dependence between the network state and the network degree. Since the average degrees of all subgraphs are similar with each other, the evolutionary network stable states of all subgraphs are also similar with each other.

C. Information Diffusion Dataset Evaluation

In this subsection, we further verify our diffusion analysis using real-world information spreading dataset [21]. In the previous two subsections, we first set the users' payoff matrix, and then conducted experiments to find the final stable state. In this experiment, an inverse step is conducted, we first estimate the evolutionary stable state through mining the real-world information spreading dataset, and then train the corresponding payoff matrix. The dataset used for experiment, named “MemeTracker”, contains more than 172 million news articles and blog posts from 1 million online sources [43]. When a site publishes a new post, it will put hyperlinks to related posts in some other sites published earlier as its sources. And later, the site will also be cited by other newer posts as well. In such a case, the hyperlinks between articles
Each group is regarded as a complete piece of information. We extract 5 groups of sites, where each phrase cluster can be regarded as a diffusion process of one piece of information. An example of one phrase cluster in the dataset is shown in Fig. 11. In such a case, each phrase cluster can be regarded as a diffusion process of one piece of information. We extract 5 groups of sites, where each group includes 500 sites. Each group is regarded as a complete graph and each site is considered as a user in our information diffusion game. We divide the dataset into two halves, where the first half is used to train the payoff matrix and the second half is used for testing. Through calculating the average hyperlinks of all phrases clusters in each group, we can obtain the statistical evolutionary stable state of each group using the first half dataset, as shown in Table II.

We can see that the sites in group 5 share major common interests, while the sites in group 1 share relatively rare common interests. Using our proposed game theoretic analysis and the data-mining based approach, enterprises/politicians can classify users into different same-interest categories according to the stable states of different information diffusions, which can help them to achieve more effective advertisement/advocacy.

Since the site network of each group is considered as a uniform degree network, by taking Table II back to Theorem 1 and normalize \( u_{fn} \) as 1, we can find the relationship between \( u_{ff} \) and \( u_{nn} \). Based on the relationship, we can parameterize the payoff matrix of each group, simulate their information diffusion processes on uniform degree graph and achieve the simulated stable information diffusion state of each group. Moreover, we can also test the stable information diffusion state of each group using the second half of the Memetracker dataset. Fig. 12 shows the simulated results from our game theoretical analysis and the estimated results from the real-world dataset, from which we can see they match well with each other. We also depict the variances of the estimated results in Fig. 12 which shows that the simulated results are always in the variance interval of the corresponding estimated results. This further verifies that our graphical evolutionary game theoretical analysis for the information diffusion over social network is effective and practical.

VI. Conclusion

In this paper, we formulate the information diffusion problem using graphical evolutionary game theory. We defined the players, strategies and payoff matrix in this problem, and highlight the correspondence between graphical EGT and information diffusion. Two kinds of networks, uniform and non-uniform degree social networks, are analyzed with the derivation of closed-form expressions of the stable network diffusion states. Moreover, we analyzed the Erdős-Rényi random network and the Barabási-Albert scale-free network. To validate our theoretical analysis, we conducted experiments on synthetic networks, real-world Facebook networks, as well as the real-world information diffusion dataset of Twitter and Memetracker. All the experiment results were consistent with the corresponding theoretical results, which corroborated that our proposed graphical EGT framework is effective and practical for modeling the information diffusion problem.
