Graphical Evolutionary Game for Information Diffusion Over Social Networks

Chunxiao Jiang, *Member, IEEE*, Yan Chen, *Member, IEEE*, and K. J. Ray Liu, *Fellow, IEEE*

Abstract—Social networks have become ubiquitous in our daily life, as such they have 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 Memetracker. Experiments shows that the proposed game theoretic framework is effective and practical in modeling the social network users' information forwarding behaviors.

Index Terms—Evolutionary game, game theory, information diffusion, information spreading, social networks.

I. INTRODUCTION

DURING the past decade, with the rapid development of the Internet and mobile technologies, social networks have become ubiquitous. Due to its diverse implication, researchers of different disciplines have been working on social networking from various perspectives [1]. A social network is a network made of connections and interactions among social entities, which can be individuals, websites, organizations and

C. Jiang was with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: chx.jiang@gmail.com).

Y. Chen and K. J. R. Liu are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: yan@umd.edu; kjrliu@umd.edu).

even intelligent equipments. Typical social network examples include Facebook/Twitter networks, hyperlink networks of websites, scientific collaboration/citation networks, and Internet of Things (IoT) [2].

Today's social networks are of very large-scale and becoming tremendous-scale, e.g., Facebook has nearly one billion active users as of September 2012 [3]. Over such large-scale social networks, users are exchanging their information, status and ideas in the form of "word-of-mouth" communications everyday. Under such a circumstance, how the information diffuses over the underlying social network has become an important research problem. On one hand, the study of information diffusion can help the enterprises/politicians to identify the influential users and links, on which the initial advertisement/advocation will be most effective. On the other hand, from the security point of view, the study of information diffusion can also help to prevent the detrimental information spreading, e.g. computer virus. The recent developments of information and communication technologies enable us to collect, store, and access the real-world big data, making the information diffusion research versatile and meaningful, at the same time more challenging.

The study of information diffusion originates from the research of computer virus/epidemic spreading over network [4]. One of the earliest and prominent works about information diffusion is [5], which studied the dynamics of information propagation through blogspace from both macroscopic and microscopic points of views. Subsequently, there are numerous works on the information diffusion, and here we summarize the works that are most representative and relevant to our study. The existing works about information diffusion can be classified into two categories. The first category focuses on analyzing the characteristics of information diffusion [6]–[10]. In [6], the authors discussed how to extract the most influential nodes on a large-scale social network. Later, a community-based greedy algorithm was proposed for mining top-k influential nodes in mobile social networks [7]. How to restrain the private or contaminated information diffusion was studied in [8] and [9] through identifying the important information links and hubs, respectively. On the other hand, how to maximize information diffusion through a network was discussed in [10] by designing effective neighbors selection strategies.

The second category focuses on analyzing the dynamic diffusion process over different kinds of networks using different mathematical models [11]–[16]. In [11], the authors studied how social networks affect the spread of behavior and investigated the effects of network structure on users' behavior diffusion. Rather than focusing on the behavior diffusion, Bakshy *et al.*

studied the role of social networks in general information diffusion through an experimental approach [12]. 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 [13], modeling the global influence of a node on the rate of diffusion on Twitter [14], illustrating the statistical mechanics of rumor spreading on Facebook [15]. Moreover, the information diffusion on overlaying social-physical networks was analyzed in [16]. 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 the roles of users' actions and decision making have not been well investigated. 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 on it, etc. This kind of interaction can often be modeled by using game theory [17]. 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 [18]. 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 to user the graphical evolutionary game to model and analyze the social network users' information forwarding strategies and the dynamic diffusion process. In the formulation, the social network topology can be treated as a graph structure and the user with new information can be regarded as the mutant. By considering the information diffusion process as the mutant spreading process, the graphical evolutionary game provides us with an ideal tool to find the final information diffusion state. Although there exists several works on using game theory to model social network [19]–[21], e.g., the authors in [19] adopted coordination game to model the epidemic diffusion and the authors in [20] studied a two-user transmission and feedback model, both the dynamic information diffusion process and the final diffusion state have not been investigated.

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 [22] 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 [18], instead, imagines that a game is played over and over again by biologically or socially conditioned players who are randomly drawn from large populations [23]. 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 [24], [25], including congestion control [26], cooperative sensing [27], cooperative peer-to-peer (P2P) streaming [28] and dynamic spectrum access [29], [30], adaptive networks [31] and also image processing area [32]. 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 $\mathcal{X} = \{1, 2, \ldots, m\}$ and payoff matrix, $\mathbf{U}$, 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 $\mathbf{p} = [p_1, \ldots, p_i, \ldots, p_m]^T$, 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 $\mathbf{q} = [q_1, q_2, \ldots, q_m]'$ against a population in state $\mathbf{p}$ is $\mathbf{q}'\mathbf{U}\mathbf{p}$. A state $\mathbf{p}^*$ is an ESS, if and only if for all different states $\mathbf{q} \neq \mathbf{p}$, $\mathbf{p}^*$ satisfies follows [33]:

$$1)\ \mathbf{q}'\mathbf{U}\mathbf{p}^* \leq \mathbf{p}^{*\prime}\mathbf{U}\mathbf{p}^*, \tag{1}$$

$$2)\ \text{if } \mathbf{q}'\mathbf{U}\mathbf{p}^* = \mathbf{p}^{*\prime}\mathbf{U}\mathbf{p}^*, \quad \mathbf{p}^{*\prime}\mathbf{U}\mathbf{q} > \mathbf{q}'\mathbf{U}\mathbf{q}. \tag{2}$$

The first condition guarantees that the average mean payoff of the population in a different state $\mathbf{q}$ against the ESS $\mathbf{p}^*$ does not exceed the average payoff of the population in $\mathbf{p}^*$, 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 $\mathbf{q}$ against itself is lower than that of state $\mathbf{p}^*$ against $\mathbf{q}$, which further guarantees the stability of the ESS $\mathbf{p}^*$. 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 $\dot{\mathbf{p}} = d\mathbf{p}/dt$ [23], i.e,

$$\mathbf{p}^* = \arg_{\mathbf{p}}(\dot{\mathbf{p}} = 0). \tag{3}$$

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 [34]. In graphical EGT, the "fitness" of a player is locally determined from interactions with all adjacent players, which is defined as [35]

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

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 α represents the selection intensity, i.e., the relative contribution of the game to fitness. The case $\alpha \rightarrow 0$ represents the limit of weak selection [36], 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 [37].

1) 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. 1(a).

2) 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. 1(b).

3) IM update rule: each player either adopts the strategy of one neighbor or remains with his/her current strategy, with
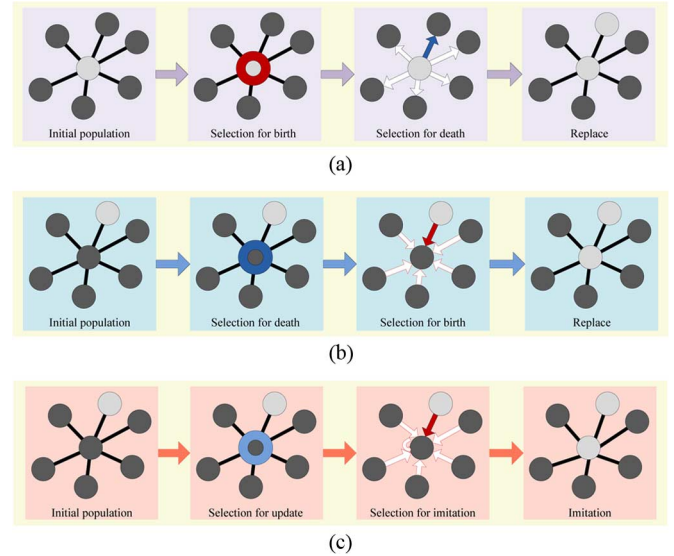

Fig. 1. Three different update rules, where death selections are shown in dark blue and birth selections are shown in red. (a) BD update rule.; (b) DB update rule.; (c) IM update rule.

the probability being proportional to fitness, as shown in Fig. 1(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. For example, in the BD update rule, a user's strategy can influence one of his/her neighbors when the fitness of adopting this strategy is high. Similarly, in the information diffusion process, when forwarding the information can bring a user more utility, the user's neighbors may also be influenced to forward the information in the near

future. Therefore, such a information diffusion process can be well modeled by the aforementioned strategy update rules.

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

$$
\begin{cases} \boldsymbol{S}_f, & \text{forward the information,} \\ \boldsymbol{S}_n, & \text{not forward the information.} \end{cases} \tag{5}
$$

Meanwhile, users' payoff matrix can be written as

$$
\begin{array}{c} \quad \boldsymbol{S}_f \quad \boldsymbol{S}_n \\ \begin{array}{c} \boldsymbol{S}_f \\ \boldsymbol{S}_n \end{array} \begin{pmatrix} u_{ff} & u_{fn} \\ u_{fn} & u_{nn} \end{pmatrix} \end{array} \tag{6}
$$

where a symmetric payoff structure is considered, i.e., when a user with strategy $\boldsymbol{S}_f$ meets a user with strategy $\boldsymbol{S}_n$, 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 $u_{nn} \geq u_{fn} \geq u_{ff}$. 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 $u_{fn} \geq u_{ff} \geq u_{nn}$. 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 [38].

Based on the above definitions of strategy and payoff function, we can further define the fitness of users with strategy $\boldsymbol{S}_f$ and strategy $\boldsymbol{S}_n$, respectively. According to (4), for a user with strategy $\boldsymbol{S}_f$ and k neighbors, among which k_f users adopting the same strategy $\boldsymbol{S}_f$ and $k - k_f$ users adopting strategy $\boldsymbol{S}_n$, his/her fitness can be written as

$$
\pi_f(k, k_f) = (1 - \alpha) + \alpha \left[k_f u_{ff} + (k - k_f) u_{fn} \right], \tag{7}
$$

where the baseline fitness is normalized as 1. Similarly, for a user with strategy $\boldsymbol{S}_n$ and k neighbors, among which k_n users adopting the same strategy $\boldsymbol{S}_n$ and $k - k_n$ users adopting strategy $\boldsymbol{S}_f$, his/her fitness can be written as

$$
\pi_n(k, k_n) = (1 - \alpha) + \alpha \left[k_n u_{nn} + (k - k_n) u_{fn} \right]. \tag{8}
$$

Table I summarizes the correspondence between the terminologies in graphical EGT and those in the social network. Based on the definitions of strategy and fitness, we can analyze the stable information diffusion state in a social network using graphical EGT under different strategy update rules. In the following two sections, we will study two scenarios: information

Graphical EGT	Social Network
Graph structure	Social network topology
Players	Users in the social network
Strategy	S_f: forward the information S_n: not forward the information
Fitness	Utility from forwarding or not
ESS	Stable information diffusion state

diffusion over uniform degree network and non-uniform degree network, respectively.

III. INFORMATION DIFFUSION OVER UNIFORM DEGREE NETWORK

In this section, we study the information diffusion process over networks with uniform degree using graphical EGT. The main reason that we discuss the uniform degree network is to provide more insight into this complicated problem, and the derivation and results (e.g., the fitness calculation and dynamics derivations) of the uniform degree network analysis will be used in the non-uniform degree case. 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 ($\boldsymbol{S}_f$ or $\boldsymbol{S}_n$). 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 $\boldsymbol{S}_f$ and $\boldsymbol{S}_n$ 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 $\boldsymbol{S}_f$ and $\boldsymbol{S}_n$, respectively; p_{fn} represents the percentage of edges on which one user chooses to forward the information while the other chooses not. Moreover, let $p_{f|n}$ denote the percentage of a user's neighbors using strategy $\boldsymbol{S}_f$, given the user is using strategy $\boldsymbol{S}_n$. Similarly, we have $p_{f|f}$, $p_{n|f}$ and $p_{n|n}$. In such a case, according to the basic probability theory, we can summarize the relationship of these defined variables as follows:

$$
p_f + p_n = 1, \tag{9}
$$

$$
p_{f|f} + p_{n|f} = 1, \ p_{f|n} + p_{n|n} = 1, \ p_f p_{n|f} = p_n p_{f|n}, \tag{10}
$$

$$
p_{ff} = p_f p_{f|f}, \ p_{fn} = p_f p_{n|f} + p_n p_{f|n}, \ p_{nn} = p_n p_{n|n}. \tag{11}
$$

Eqns. (9) and (11) imply that the state of the whole social network can be described by only two variables, p_f and p_{ff}. In
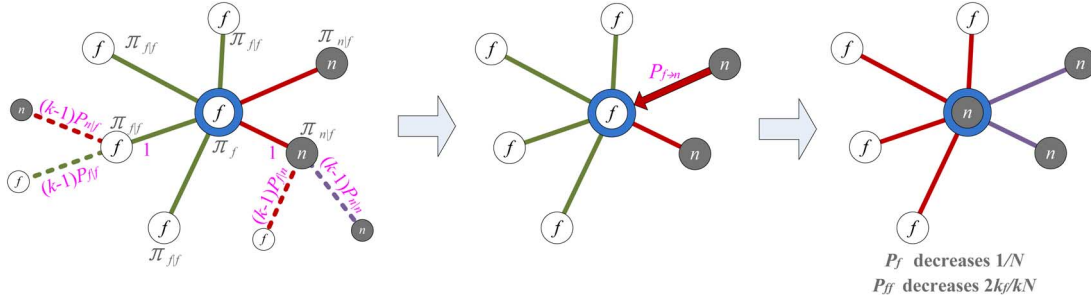
Fig. 2. Evolution of user with strategy $\boldsymbol{S}_f$, where f means using strategy $\boldsymbol{S}_f$ and n means using strategy $\boldsymbol{S}_n$.

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. 2 illustrates a scenario where the center user adopts strategy $\boldsymbol{S}_f$, and among his/her k neighbors, there are k_f users adopting the same strategy $\boldsymbol{S}_f$ and $k - k_f$ users adopting the opposite strategy $\boldsymbol{S}_n$. The probability of such a configuration is

$$\theta_f(k, k_f) = \binom{k}{k_f} p_{f|f}^{k_f} (1 - p_{f|f})^{k-k_f}. \quad (12)$$

According to the fitness definition in (7), the fitness of the center user in the left part of Fig. 2, who adopts strategy $\boldsymbol{S}_f$, can be calculated by

$$\pi_f = (1 - \alpha) + \alpha \left[k_f u_{ff} + (k - k_f) u_{fn} \right]. \quad (13)$$

Among the center user's k neighbors, for the neighbors adopting strategy $\boldsymbol{S}_f$, each of them has $(k - 1)p_{n|f}$ neighbors using strategy $\boldsymbol{S}_n$ and $(k - 1)p_{f|f} + 1$ neighbors using strategy $\boldsymbol{S}_f$ averagely, where "1" means the center user with strategy $\boldsymbol{S}_f$, as shown by the dashed lines in the left part of Fig. 2. In such a case, the average fitness of the center user's neighbors who adopt strategy $\boldsymbol{S}_f$, i.e., forward the information, is

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

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

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

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. 2, the probability that the center user updates his/her current strategy $\boldsymbol{S}_f$ by strategy $\boldsymbol{S}_n$ is

$$P_{f \to n}(k_f) = \frac{(k - k_f)\pi_{n|f}}{k_f \pi_{f|f} + (k - k_f)\pi_{n|f} + \pi_f}. \quad (16)$$

Therefore, as shown in right part of Fig. 2, the center user deviating from using strategy $\boldsymbol{S}_f$ leads to the percentage of users adopting strategy $\boldsymbol{S}_f$, p_f, decreasing by $1/N$ with probability

$$\text{Prob}\left(\Delta p_f = -\frac{1}{N}\right) = \sum_{k_f=0}^{k} p_f \theta_f(k, k_f) P_{f \to n}(k_f)$$

$$= p_f \sum_{k_f=0}^{k} \binom{k}{k_f} p_{f|f}^{k_f} (1 - p_{f|f})^{k-k_f}$$

$$\cdot \frac{(k - k_f)\pi_{n|f}}{k_f \pi_{f|f} + (k - k_f)\pi_{n|f} + \pi_f}. \quad (17)$$

Meanwhile, the center user deviating from using strategy $\boldsymbol{S}_f$ also leads to the edges on which both users adopting strategy $\boldsymbol{S}_f$ decreasing by k_f, which means that p_{ff} decreases by $2k_f/kN$ with probability

$$\text{Prob}\left(\Delta p_{ff} = \frac{-2k_f}{kN}\right) = p_f \binom{k}{k_f} p_{f|f}^{k_f} (1 - p_{f|f})^{k-k_f}$$

$$\cdot \frac{(k - k_f)\pi_{n|f}}{k_f \pi_{f|f} + (k - k_f)\pi_{n|f} + \pi_f}. \quad (18)$$

Similar analysis can be applied to the user adopting strategy $\boldsymbol{S}_n$. According to the IM update rule, a user adopting strategy $\boldsymbol{S}_n$ is selected for imitation with probability $p_n = 1 - p_f$. As shown in Fig. 3, among its k neighbors, suppose that there are k_n users adopting the same strategy $\boldsymbol{S}_n$ and $k - k_n$ users adopting the opposite strategy $\boldsymbol{S}_f$. The probability of such a configuration is

$$\theta_n(k, k_n) = \binom{k}{k_n} p_{n|n}^{k_n} (1 - p_{n|n})^{k-k_n}. \quad (19)$$

Under such a configuration, the fitness of the center user who adopts strategy $\boldsymbol{S}_n$ is

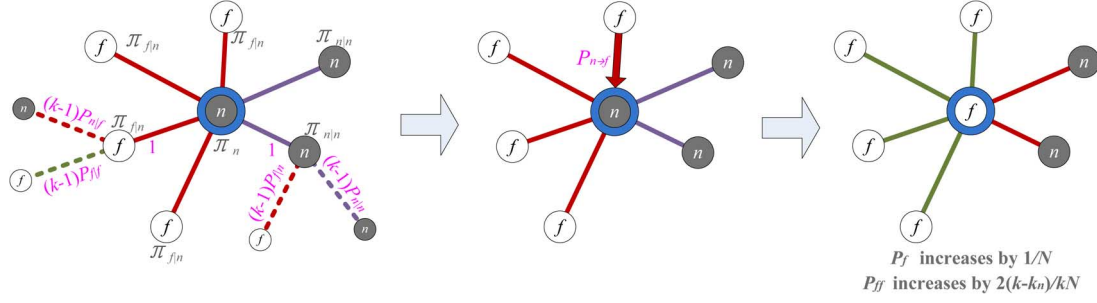$$\pi_n = (1 - \alpha) + \alpha \left[k_n u_{nn} + (k - k_n) u_{fn} \right]. \quad (20)$$

Fig. 3. Evolution of user with strategy $\boldsymbol{S}_n$, where f means using strategy $\boldsymbol{S}_f$ and n means using strategy $\boldsymbol{S}_n$.

Among his/her k neighbors, the average fitness of neighbors who adopt strategy $\boldsymbol{S}_f$ is

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

and the fitness of node adopting strategy $\boldsymbol{S}_n$ is

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

In such a case, as shown in the middle part of Fig. 3, the probability that the center user updates his/her strategy $\boldsymbol{S}_n$ by strategy $\boldsymbol{S}_f$ is

$$P_{n\to f}(k_n) = \frac{(k-k_n)\pi_{f|n}}{k_n\pi_{n|n} + (k-k_n)\pi_{f|n} + \pi_n}. \tag{23}$$

Therefore, as shown in right part of Fig. 3, the center user deviating from using strategy $\boldsymbol{S}_n$ leads to the percentage of users adopting strategy $\boldsymbol{S}_f$, p_f, increasing by $1/N$ with probability

$$\mathrm{Prob}\left(\Delta p_f = \frac{1}{N} \right) = \sum_{k_n=0}^{k} (1-p_f)\theta_n(k,k_n)P_{n\to f}(k_n)$$

$$= (1-p_f) \sum_{k_n=0}^{k} \binom{k}{k_n} p_{n|n}^{k_n}(1-p_{n|n})^{k-k_n}$$

$$\cdot \frac{(k-k_n)\pi_{f|n}}{k_n\pi_{n|n} + (k-k_n)\pi_{f|n} + \pi_n}. \tag{24}$$

Meanwhile, the deviation also leads to the edges on which both users adopting strategy $\boldsymbol{S}_f$ increasing by $k - k_n$, thus, we have

$$\mathrm{Prob}\left(\Delta p_{ff} = \frac{k-k_n}{\frac{kN}{2}} \right) = (1-p_f)\binom{k}{k_n} p_{n|n}^{k_n}(1-p_{n|n})^{k-k_n}$$

$$\cdot \frac{(k-k_n)\pi_{f|n}}{k_n\pi_{n|n} + (k-k_n)\pi_{f|n} + \pi_n}. \tag{25}$$

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

$$\dot{p}_f = \frac{1}{N}\mathrm{Prob}\left(\Delta p_f = \frac{1}{N} \right) - \frac{1}{N}\mathrm{Prob}\left(\Delta p_f = -\frac{1}{N} \right)$$

$$= \frac{\alpha k(k-1)p_{fn}}{2N(k+1)^2}(\gamma_1 u_{nn} + \gamma_2 u_{fn} + \gamma_3 u_{ff}) + O(\alpha^2), \tag{26}$$

where the second equality is according to Taylor's Theorem and weak selection assumption with α goes to zero [39], and the parameters γ_1, γ_2 and γ_3 are given as follows:

$$\gamma_1 = -p_{n|n}\left[(k-1)(p_{n|n} + p_{f|f}) + 3 \right], \tag{27}$$

$$\gamma_2 = p_{n|n} - p_{f|f} + (p_{f|n} - p_{f|n})$$
$$\cdot \left[(k-1)(p_{n|n} + p_{f|f}) + 2 \right], \tag{28}$$

$$\gamma_3 = p_{f|f}\left[(k-1)(p_{n|n} + p_{f|f}) + 3 \right]. \tag{29}$$

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 [36]. Moreover, the weak selection has a long tradition in theoretical biology [40]. 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

$$\dot{p}_{ff} = \sum_{k_f=0}^{k} \frac{-2k_f}{kN}\mathrm{Prob}\left(\Delta p_{ff} = \frac{-2k_f}{kN} \right)$$

$$+ \sum_{k_n=0}^{k} \frac{2(k-k_n)}{kN}\mathrm{Prob}\left(\Delta p_{ff} = \frac{2(k-k_n)}{kN} \right)$$

$$= \frac{p_{fn}}{(k+1)N}\left(1 + (k-1)(p_{f|n} - p_{f|f}) \right) + O(\alpha). \tag{30}$$

B. ESS of Information Diffusion

Based on the evolution dynamics analysis, we can obtain the evolutionary stable network state by jointly solving (26) and (30), as shown in the following theorem.

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:

$$p_f^* = \begin{cases} 1, & \text{if } u_{ff} > u_{fn} > u_{nn}, \\ 0, & \text{if } u_{nn} > u_{fn} > u_{ff}, \\ \frac{(k-2)u_{fn}+u_{ff}-(k-1)u_{nn}}{2u_{fn}-(k-2)(u_{ff}-u_{nn})}, & \text{else.} \end{cases} \quad (31)$$

Proof: See Appendix. ∎

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 to *Theorem 1*, we can see that the condition of $p_f^* = 1$ means that both users forwarding the information can gain the most payoff, while not forwarding gains the least payoff. In a social network, this is corresponding to the scenario where the released information is an extremely hot topic, forwarding which can attract more attentions. On the contrary, the condition of $p_f^* = 0$ means that neither user forwarding the information can gain the most payoff, while both forwarding gains the least payoff, which is corresponding to the scenario where the released information is useless or negative advertisement, forwarding which can only incur unnecessary cost. For the third ESS, $p_f^* = b/a$, some approximations can be made as follows:

$$\begin{aligned} p_f^* = \frac{b}{a} &= \frac{(k-2)u_{fn}+u_{ff}-(k-1)u_{nn}}{(k-2)(2u_{fn}-u_{ff}-u_{nn})} \\ &= \frac{(k-1)(u_{fn}-u_{nn})+(u_{ff}-u_{fn})}{(k-2)(u_{fn}-u_{ff}+u_{fn}-u_{nn})} \\ &\doteq \frac{1}{1+\frac{u_{fn}-u_{ff}}{u_{fn}-u_{nn}}}, \end{aligned} \quad (32)$$

where the last approximation is based on the assumption that the network degree $k \gg 2$ in real social networks. We can see that when $k \gg 2$, the information diffusion result in uniform degree network shows the scale-free phenomenon. Meanwhile, we can further discuss some physical meanings of the third ESS as follows.

• *Case 1*: $u_{fn} > u_{ff}$ and $u_{fn} > u_{nn}$.

In this case, unilateral forwarding can bring more payoff than no forwarding or both forwarding. In a social network, this case is corresponding to the scenario where both users forwarding the information can gain only limited reward but incurring more cost to both of them. An example of this case can be that the information is not the mainstream topic, e.g. the news about a punk musician, and is supposed to be diffused among people with similar interstates. Specifically, when $u_{ff} = u_{nn}$, we have $p_f^* \approx 0.5$. This means that if there is no payoff difference between forwarding and not forwarding the information, the final stable network state tends to "half and half". Moreover, when $u_{fn} > u_{ff} > u_{nn}$, we have $p_f^* > 0.5$, i.e., more than half of users' strategies are forwarding; while when $u_{fn} > u_{nn} > u_{ff}$, we have $p_f^* < 0.5$, i.e., less than half of users' forward the information.

• *Case 2*: $u_{ff} > u_{fn}$ and $u_{nn} > u_{fn}$.

This is the opposite case of *Case 1*. In this case, the payoff configuration is equivalent to that of the coordination game, where both players with the same actions can make more payoff than opposite actions. In a social network, this case is corresponding to the scenario where both users forwarding the information can enhance the popularity of them simultaneously

with some forwarding costs. However, if only one user forwards the information, the associated cost is larger than the unilateral benefit. An example of this case can be that the information is politically sensitive and the reality of it is not guaranteed, forwarding which may gain attractions but also incurring potential misleading cost.

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

$$\overline{k} = \sum_{k=0}^{+\infty} \lambda(k)k. \quad (33)$$

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, (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 [41]. 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. 2, 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

$$P_{n \to f}(l, l_f) = p_f \theta_f(l, l_f) \frac{\pi_f(l, l_f)}{\overline{\pi}} \frac{l - l_f}{l}, \quad (34)$$

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

$\overline{\pi}$ is the expected fitness of all users which can be calculated as follows:

$$\overline{\pi} = N p_f \sum_{k=0}^{+\infty} \lambda(k) \left(k p_{f|f} u_{ff} + k(1 - p_{f|f}) u_{fn} \right)$$
$$+ N p_n \sum_{k=0}^{+\infty} \lambda(k) \left(k p_{n|n} u_{nn} + k(1 - p_{n|n}) u_{fn} \right)$$
$$= N(p_{ff} u_{ff} + p_{fn} u_{fn} + p_{nn} u_{nn}). \quad (35)$$

Therefore, the percentage of users adopting strategy $\boldsymbol{S}_f$, p_f, increases by $1/N$ with expected probability

$$\text{Prob}\left(\Delta p_f = \frac{1}{N} \right) = \sum_{l=0}^{+\infty} \lambda(l) \sum_{l_f=0}^{l} p_f \theta_f(l, l_f) \frac{\pi_f(l, l_f)}{\overline{\pi}} \frac{l - l_f}{l}. \quad (36)$$

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

$$\text{Prob}\left(\Delta p_{ff} = \frac{(l' - 1)p_{f|n} + 1}{\frac{\overline{k} N}{2}} \right) = \text{Prob}\left(\Delta p_f = \frac{1}{N} \right), \quad (37)$$

where $\overline{k}$ is the average degree of the network given in (33).

Similarly, suppose the chosen user is adopting strategy $\boldsymbol{S}_n$ and with degree m as shown in Fig. 3. Among his/her neighbors, there are m_n users adopting the same strategy $\boldsymbol{S}_f$ and $m - m_n$ users adopting the opposite strategy $\boldsymbol{S}_n$. In such a case, according to the BD update rule, the percentage of users adopting strategy $\boldsymbol{S}_f$, p_f, decreases by $1/N$ with expected probability

$$\text{Prob}\left(\Delta p_f = -\frac{1}{N} \right)$$
$$= \sum_{m=0}^{+\infty} \lambda(m) \sum_{m_f=0}^{m} (1 - p_f) \theta_n(m, m_f) \frac{\pi_n(m, m_f)}{\overline{\pi}} \frac{m_f}{m}, \quad (38)$$

where $\theta_n(m, m_f)$ is given in (19), $\pi_n(m, m_f)$ is given in (8) and $\overline{\pi}$ is given in (35). Suppose the replaced user, who was adopting strategy $\boldsymbol{S}_f$, is with degree m'. After strategy update, the edges on which both users adopting strategy $\boldsymbol{S}_f$ will decrease by $(m' - 1)p_{f|f}$. In such a case, we have

$$\text{Prob}\left(\Delta p_{ff} = -\frac{(m' - 1)p_{f|f}}{\frac{\overline{k} N}{2}} \right) = \text{Prob}\left(\Delta p_f = -\frac{1}{N} \right). \quad (39)$$

Combining (36) and (38), we have the dynamics of p_f as

$$\dot{p}_f = \frac{1}{N} \text{Prob}\left(\Delta p_f = \frac{1}{N} \right) - \frac{1}{N} \text{Prob}\left(\Delta p_f = -\frac{1}{N} \right)$$
$$= \frac{\alpha(\overline{k} - 1)p_{fn}}{2\overline{\pi} N} \left[p_{f|f}(u_{ff} - u_{fn}) - p_{n|n}(u_{nn} - u_{fn}) \right]$$
$$+ O(\alpha^2), \quad (40)$$

where $\overline{k}$ is the average degree of the network given in (33). Combining (37) and (39), we have the dynamics of p_{ff} as

$$\dot{p}_{ff} = \frac{(\overline{l'} - 1)p_{f|n} + 1}{\frac{\overline{k} N}{2}} \cdot \text{Prob}\left(\Delta p_f = \frac{1}{N} \right)$$
$$- \frac{(\overline{m'} - 1)p_{f|f}}{\frac{\overline{k} N}{2}} \cdot \text{Prob}\left(\Delta p_f = -\frac{1}{N} \right), \quad (41)$$

where $\overline{l'}$ and $\overline{m'}$ are the average degrees of the replaced users during the strategy update, which are different with the average degree of the whole network, $\overline{k}$. If a link is selected at random, the degree distribution of the node on the specific link is not $\lambda(k)$ but rather $k\lambda(k)/\sum_{k=0}^{+\infty} k\lambda(k)$ [42]. Thus, the average degree of the replaced users is

$$\overline{l'} = \overline{m'} = \sum_{k=0}^{+\infty} k \frac{k\lambda(k)}{\sum_{k=0}^{+\infty} k\lambda(k)} = \frac{\overline{k^2}}{\overline{k}}, \quad (42)$$

where $\overline{k^2} = \sum_{k=0}^{+\infty} k^2 \lambda(k)$ is the expectoration of k^2. Based on the dynamics of p_f and p_{ff} in (40) and (41), 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:

$$p_f^* = \begin{cases} 1, & \text{if } u_{ff} > u_{fn} > u_{nn}, \\ 0, & \text{if } u_{nn} > u_{fn} > u_{ff}, \\ \dfrac{\left(\frac{\overline{k^2}}{\overline{k} - 2} \right)(u_{fn} - u_{nn}) + (u_{ff} - u_{nn})}{\left(\frac{\overline{k^2}}{\overline{k} - 2} \right)(2u_{fn} - u_{ff} - u_{nn})}, & \text{else.} \end{cases} \quad (43)$$

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 (41) and (42). Meanwhile, the conditions in (43) can be obtained by analyzing the Jacob matrix of $\dot{p}_f$ and $\dot{p}_{ff}$. Due to page limitation, the detailed derivation is omitted. ∎

Since $\overline{k^2} \geq (\overline{k})^2$, i.e., $\overline{k^2}/\overline{k} \geq \overline{k}$, we can see that when $\overline{k} \gg 2$, the information diffusion result in non-uniform degree network shows the scale-free phenomenon.

B. Special Cases

In this subsection, we discuss two special cases of the non-uniform degree networks, Erdős-Rényi random network [43] and Barabási-Albert scale-free network [44]. For the Erdős-Rényi random (ER) network, the degree follows a Poisson distribution, i.e.,

$$\lambda_{\text{ER}}(k) = \frac{e^{-\overline{k}} \overline{k}^k}{k!} \text{ and } \overline{k^2} = \overline{k}(\overline{k} + 1). \quad (44)$$

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

$$p_{f\text{ER}}^* = \begin{cases} 1, & \text{if } u_{ff} > u_{fn} > u_{nn}, \\ 0, & \text{if } u_{nn} > u_{fn} > u_{ff}, \\ \dfrac{(\overline{k} - 1)(u_{fn} - u_{nn}) + (u_{ff} - u_{nn})}{(\overline{k} - 1)(2u_{fn} - u_{ff} - u_{nn})}, & \text{else.} \end{cases} \quad (45)$$

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

$$\lambda_{\mathrm{BA}}(k) \propto k^{-\xi} \text{ and } \overline{k^2} \doteq \overline{k}^2 \log N/4 \text{ (when } \xi = 3). \quad (46)$$

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

$$p_{f\mathrm{BA}}^*$$
$$= \begin{cases} 1, & \text{if } u_{ff} > u_{fn} > u_{nn}, \\ 0, & \text{if } u_{nn} > u_{fn} > u_{ff}, \\ \frac{(\overline{k}\log N - 8)(u_{fn} - u_{nn}) + 4(u_{ff} - u_{nn})}{(\overline{k}\log N - 8)(2u_{fn} - u_{ff} - u_{nn})}, & \text{else.} \end{cases} \quad (47)$$

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: $u_{ff} > u_{fn} > u_{nn}$

$$\begin{array}{cc} & \begin{array}{cc} \boldsymbol{S}_f & \boldsymbol{S}_n \end{array} \\ \begin{array}{c} \boldsymbol{S}_f \\ \boldsymbol{S}_n \end{array} & \begin{pmatrix} 0.8 & 0.6 \\ 0.6 & 0.4 \end{pmatrix} \end{array} \quad (48)$$

- PM 2: $u_{fn} > u_{ff} > u_{nn}$

$$\begin{array}{cc} & \begin{array}{cc} \boldsymbol{S}_f & \boldsymbol{S}_n \end{array} \\ \begin{array}{c} \boldsymbol{S}_f \\ \boldsymbol{S}_n \end{array} & \begin{pmatrix} 0.6 & 0.8 \\ 0.8 & 0.4 \end{pmatrix} \end{array} \quad (49)$$

- PM 3: $u_{fn} > u_{nn} > u_{ff}$

$$\begin{array}{cc} & \begin{array}{cc} \boldsymbol{S}_f & \boldsymbol{S}_n \end{array} \\ \begin{array}{c} \boldsymbol{S}_f \\ \boldsymbol{S}_n \end{array} & \begin{pmatrix} 0.4 & 0.8 \\ 0.8 & 0.6 \end{pmatrix} \end{array} \quad (50)$$

- PM 4: $u_{nn} > u_{fn} > u_{ff}$

$$\begin{array}{cc} & \begin{array}{cc} \boldsymbol{S}_f & \boldsymbol{S}_n \end{array} \\ \begin{array}{c} \boldsymbol{S}_f \\ \boldsymbol{S}_n \end{array} & \begin{pmatrix} 0.4 & 0.6 \\ 0.6 & 0.8 \end{pmatrix} \end{array} \quad (51)$$

Fig. 4 shows the experiment results for the uniform-degree network under different average degrees and payoff matrices.
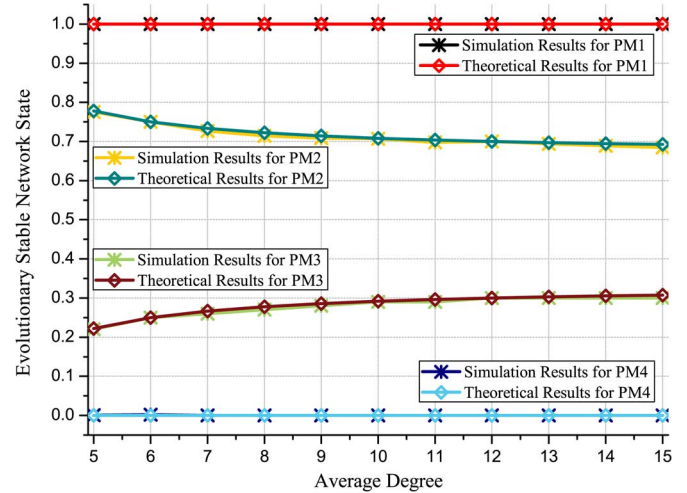

Fig. 4. Simulation results for the uniform-degree network.

The theoretical results are calculated from (31) 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. 4, 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 (32).

Fig. 5 and 6 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 (45) and (47) 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. 6, 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 [45]. The Facebook graph dataset contains 4039 users and 88234 edges, where the edge represents the connection between two users [46]. The graph structure of the Facebook network is depicted in Fig. 7, where the nodes with large degrees, i.e., the users with large number of friends, are highlighted. Meanwhile,
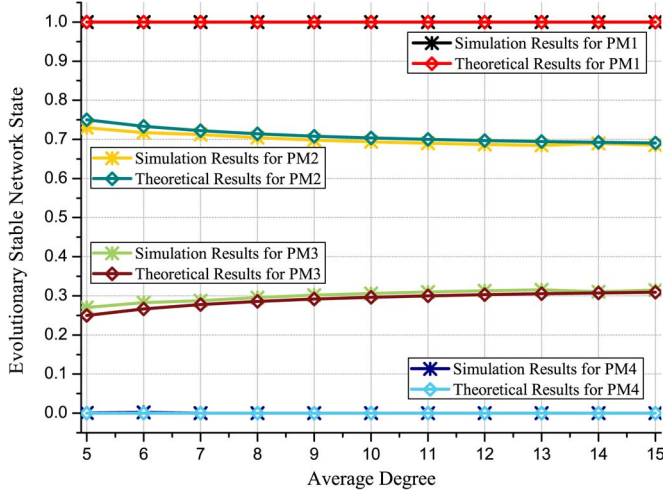
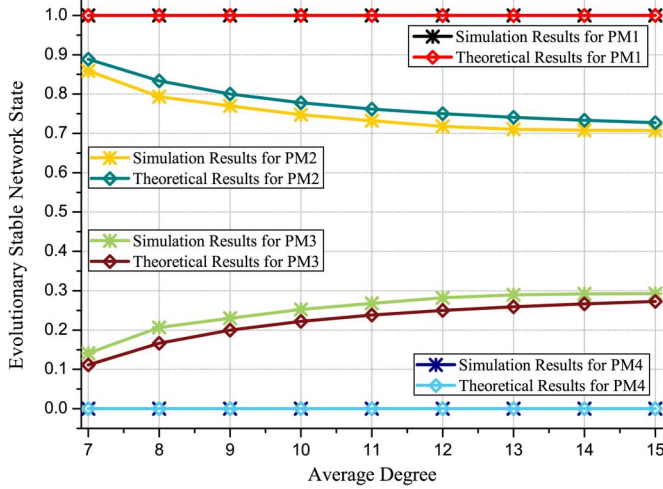Fig. 5. Simulation results for the Erdős-Rényi random network.

Fig. 6. Simulation results for the Barabási-Albert scale-free network.
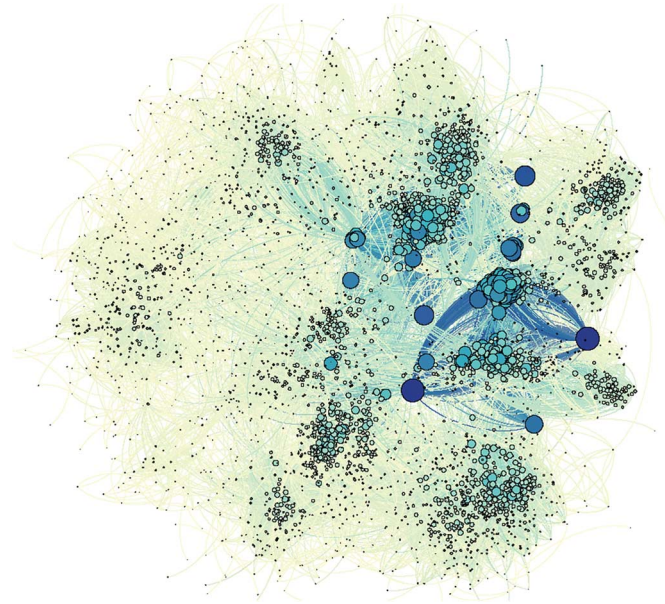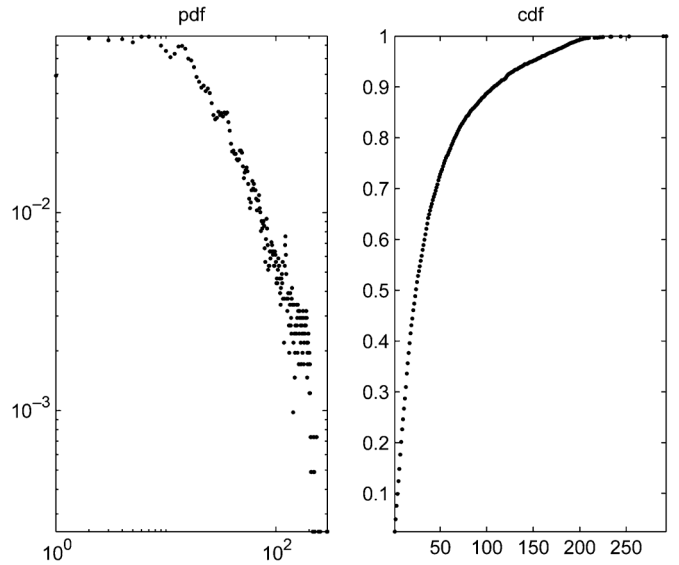

Fig. 7. Graph structure of the Facebook network used for experiment.

Fig. 8. Degree characteristics of the Facebook network.

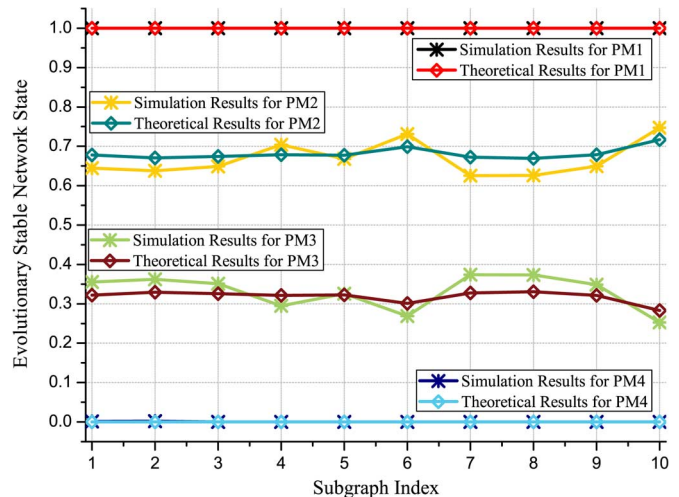Fig. 9. Experiment result for the Facebook network.

we also plot the degree characteristics of the Facebook network in Fig. 8, 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 axis are in logarithm scale, we can see that the Facebook network degree exhibits the scale-free phenomenon [44].

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. 9 shows the experiment results under different payoff matrix settings. The theoretical results are calculated from (43), 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.

```
2  8  we're not commenting on that story i'm afraid    2131865
    3  3  we're not commenting on that            489007
       2008-08-18 14:23:05  1  M  http://business.theage.com.au/business/bb-chief-set-to-walk-plank-20080818-3xp7.html
       2008-11-26 01:27:13  1  B  http://sfweekly.com/2008-11-26/news/buy-line
       2008-11-27 18:55:30  1  B  http://aconstantineblacklist.blogspot.com/2008/11/re-researcher-matt-janovic.html
    5  2  we're not commenting on that story        2131864
       2008-12-08 14:50:18  3  B  http://videogaming247.com/2008/12/08/home-in-10-days-were-not-commenting-on-that-story-says-scee
       2008-12-08 19:35:31  2  B  http://jplaystation.com/2008/12/08/home-in-10-days-were-not-commenting-on-that-story-says-scee
```

Fig. 10.   Example of the Memetracker phrase cluster dataset [22].

TABLE II
EXPERIMENT RESULTS FOR MEMETRACKER DATASET

Group Index	1	2	3	4	5
ESS	0.19	0.35	0.53	0.77	0.81

C. Information Diffusion Dataset Evaluation

In this subsection, we further verify our diffusion analysis using real-world information spreading dataset [22]. 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, where 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 "Meme-Tracker," contains more than 172 million news articles and blog posts from 1 million online sources [47]. 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 and posts can be used to represent the spreading of information from one site to another site.

In the MemeTracker phrase cluster dataset, each "phrase cluster" contains all the plain and mutant phrases in one cluster, a list of sites where the phrases appeared, and time indexes when the phrase appeared. An example of one phrase cluster in the dataset is shown in Fig. 10. In such a case, each phrase cluster can be regarded as a diffusion process of one piece of information. We extract 5 group 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/advocation.

Since the site network of each group is considered as a uniform degree network, by taking Table II back to *Theorem 1* and normalizing u_{fn} as 1, we can find the relationship between u_{ff} and u_{nn}. Based on the relationship, we can parameterize the
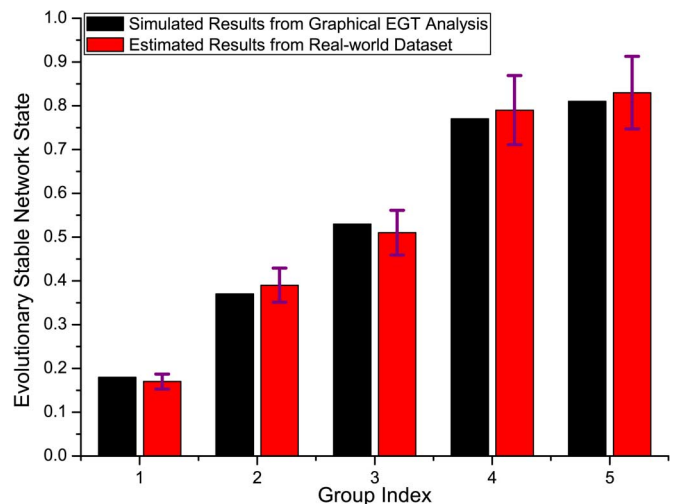

Fig. 11.   Experiment result for Memetracker dataset.

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. Meanwhile, we can also test the stable information diffusion state of each group using the second half of the Memetracker dataset. Fig. 11 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. 11, 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 we real-world information diffusion dataset of Memetracker. All the experiment results were consistent with corresponding theoretical results, which corroborated that our proposed graphical EGT framework is effective and practical for modeling the information diffusion problem.

APPENDIX
PROOF OF THEOREM 1

Proof: 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 $\dot{p}_{ff} = 0$ in (30), we have

$$p_{f|f} - p_{f|n} = \frac{1}{k-1}. \tag{52}$$

In such a case, according to (9)–(11) and (52), all the network state parameters can be expressed only by p_f as follow:

$$p_{f|f} = p_f + \frac{1}{k-1}(1 - p_f), \qquad (53)$$

$$p_{f|n} = \frac{k-2}{k-1}p_f, \qquad (54)$$

$$p_{n|f} = \frac{k-2}{k-1}(1 - p_f), \qquad (55)$$

$$p_{n|n} = 1 - \frac{k-2}{k-1}p_f. \qquad (56)$$

Then, by substituting (53)–(56) and (27)–(29) into (26), we can obtain the dynamic of the network user state, $\dot{p}_f$, as

$$\dot{p}_f = \frac{\alpha k(k-2)(k+3)}{2N(k-1)(k+1)^2}p_f(1-p_f)(ap_f - b), \qquad (57)$$

where the parameters a and b are given as follows:

$$a = (k-2)(u_{ff} - 2u_{fn} + u_{nn}), \qquad (58)$$
$$b = (k-1)u_{nn} - (k-2)u_{fn} - u_{ff}. \qquad (59)$$

At the ESS, we have $\dot{p}_f = 0$. Therefore, there are three possible evolutionary stable network states:

$$p_f^* = 0, 1 \text{ and } \frac{b}{a}. \qquad (60)$$

Considering (26) and (30) as a nonlinear dynamic system, we can examine whether the three stable points in (60) are evolutionary stable network states through analyzing the Jacobian matrix of the dynamics as follows:

$$\mathbf{J} = \begin{bmatrix} \frac{\partial \dot{p}_f}{\partial p_f} & \frac{\partial \dot{p}_f}{\partial p_{ff}} \\ \frac{\partial \dot{p}_{ff}}{\partial p_f} & \frac{\partial \dot{p}_{ff}}{\partial p_{ff}} \end{bmatrix}. \qquad (61)$$

The evolutionary stability requires that $det(\mathbf{J}) > 0$ and $tr(\mathbf{J}) < 0$ [18]. By substituting (26) and (30) into (61), we can obtain the three conditions for $p_f^* = 0$, 1 and b/a, respectively. Due to page limitation, the detailed derivations are omitted here. ∎

REFERENCES

[1] H. V. Zhao, W. S. Lin, and K. J. R. Liu, *Behavior Dynamics in Media-Sharing Social Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[2] Wikipedia, Social Network, [Online]. Available: http://en.wikipedia.org/wiki/Social_network

[3] Wikipedia, Facebook, [Online]. Available: http://en.wikipedia.org/wiki/Facebook

[4] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," *Phys. Rev. Lett.*, vol. 86, no. 14, pp. 3200–3203, 2001.

[5] D. Gruhl, R. Guha, D. Liben-Nowell, and A. Tomkins, "Information diffusion through blogspace," in *Proc. 13rd Int. Conf. World Wide Web (WWW)*, 2004, pp. 491–501.

[6] M. Kimura, K. Saito, and R. Nakano, "Extracting influential nodes for information diffusion on a social network," in *Prof. of AAAI Conf. Artif. Intell.*, 2007, pp. 1371–1376.

[7] Y. Wang, G. Cong, G. Song, and K. Xie, "Community-based greedy algorithm for mining top-k influential nodes in mobile social networks," in *Proc. 16th ACM SIGKDD Int. Conf. Knowl. Disc. Data Mining*, 2010, pp. 1039–1048.

[8] M. Kimura, K. Saito, and H. Motoda, "Blocking links to minimize contamination spread in a social network," *ACM Trans. Knowl. Discov. Data*, vol. 3, no. 2, pp. 9:1–9:23, 2009.

[9] M. U. Ilyas, M. Z. Shafiq, A. X. Liu, and H. Radha, "A distributed and privacy preserving algorithm for identifying information hubs in social networks," in *Prof. IEEE INFOCOM*, 2011, pp. 561–565.

[10] H. Kim and E. Yoneki, "Influential neighbors selection for information diffusion in online social networks," in *Prof. IEEE ICCCN*, 2012, pp. 1–7.

[11] D. Centola, "The spread of behavior in an online social network experiment," *Science*, vol. 329, no. 5996, pp. 1194–1197, 2010.

[12] E. Bakshy, I. Rosenn, C. Marlow, and L. Adamic, "The role of social networks in information diffusion," in *Proc. 21st Int. Conf. World Wide Web (WWW)*, 2012, pp. 519–528.

[13] J. Yang and S. Counts, "Predicting the speed, scale, range of information diffusion in twitter," in *Proc. 4th Int. AAAI Conf. Weblogs and Social Media*, 2010, pp. 355–358.

[14] J. Yang and J. Leskovec, "Modeling information diffusion in implicit networks," in *Proc. IEEE Int. Conf. Data Mining*, 2010, pp. 599–608.

[15] M. Ostilli, E. Yoneki, I. X. Leung, J. F. Mendes, P. Lio, and J. Crowcroft, "Statistical mechanics of rumor spreading in network communities," in *Proc. Int. Conf. Comput. Sci.*, 2010, pp. 2331–2339.

[16] O. Yagan, D. Qian, J. Zhang, and D. Cochran, "Conjoining speeds up information diffusion in overlaying social-physical networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 6, pp. 1038–1048, Jun. 2013.

[17] Y. Chen and K. J. R. Liu, "Understanding microeconomic behaviors in social networking: An engineering view," *IEEE Signal Process. Mag.*, vol. 29, no. 2, pp. 53–64, Mar. 2012.

[18] J. M. Smith, *Evolution and the Theory of Games*. Cambridge, U.K.: Cambridege Univ. Press, 1982.

[19] F. Vega-Redondo, *Complex Social Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2007.

[20] D. Zinoviev and V. Duong, "A game theoretical approach to broadcast information diffusion in social networks," in *Proc. Annu. Simulat. Symp. (ANSS)*, 2011, pp. 47–52.

[21] W. Qiu, Y. Wang, and J. Yu, "A game theoretical model of information dissemination in social network," in *Proc. Int. Conf. Complex Syst. (ICCS)*, 2012, pp. 1–6.

[22] J. Leskovec, Memetracker Data, [Online]. Available: http://www.memetracker.org

[23] R. Cressman, *Evolutionary Dynamics and Extensive Form Games*. Cambridge, MA, USA: MIT Press, 2003.

[24] K. J. R. Liu and B. Wang, *Cognitive Radio Networking and Security: A Game Theoretical View*. Cambridge, U.K.: Cambridge Univ. Press, 2010.

[25] B. Wang, Y. Wu, and K. J. R. Liu, "Game theory for cognitive radio networks: An overview," *Comput. Netw.*, vol. 54, no. 14, pp. 2537–2561, 2010.

[26] E. H. Watanabe, D. Menasché, E. Silva, and R. M. Leão, "Modeling resource sharing dynamics of VoIP users over a WLAN using a game-theoretic approach," in *Proc. IEEE INFOCOM*, 2008, pp. 915–923.

[27] B. Wang, K. J. R. Liu, and T. C. Clancy, "Evolutionary cooperative spectrum sensing game: How to collaborate?," *IEEE Trans. Commun.*, vol. 58, no. 3, pp. 890–900, 2010.

[28] Y. Chen, B. Wang, W. S. Lin, Y. Wu, and K. J. R. Liu, "Cooperative peer-to-peer streaming: An evolutionary game-theoretic approach," *IEEE Trans. Circuit Syst. Video Technol.*, vol. 20, no. 10, pp. 1346–1357, 2010.

[29] C. Jiang, Y. Chen, Y. Gao, and K. J. R. Liu, "Evolutionary game for joint spectrum sensing and access in cognitive radio networks," in *Proc. IEEE GLOBECOM*, 2013, pp. 1–5.

[30] C. Jiang, Y. Chen, Y. Gao, and K. J. R. Liu, "Joint spectrum sensing and access evolutionary game in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2470–2483, May 2013.

[31] C. Jiang, Y. Chen, and K. J. R. Liu, "Distributed adaptive filtering: A graphical evolutionary game theoretic view," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5675–5688, Nov.15 2013.

[32] Y. Chen, Y. Gao, and K. J. R. Liu, "An evolutionary game-theoretic approach for image interpolation," in *Proc. IEEE ICASSP*, 2011, pp. 989–992.

[33] I. M. Bomze and E. E. van Damme, "A dynamical characterization of evolutionarily stable states," *Ann. Operat. Res.*, vol. 37, no. 1, pp. 229–244, 1992.

[34] P. Shakarian, P. Roos, and A. Johnson, "A review of evolutionary graph theory with applications to game theory," *Biosystems*, vol. 107, no. 2, pp. 66–80, 2012.

[35] M. A. Nowak and K. Sigmund, "Evolutionary dynamics of biological games," *Science*, vol. 303, pp. 793–799, 2004.

[36] H. Ohtsuki, M. A. Nowak, and J. M. Pacheco, "Breaking the symmetry between interaction and replacement in evolutionary dynamics on graphs," *Phys. Rev. Lett.*, vol. 98, no. 10, p. 108106, 2007.

[37] H. Ohtsukia and M. A. Nowak, "The replicator equation on graphs," *J. Theor. Biol.*, vol. 243, pp. 86–97, 2006.

[38] R. Cooper, *Coordination Games*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[39] F. Fu, L. Wang, M. A. Nowak, and C. Hauert, "Evolutionary dynamics on graphs: Efficient method for weak selection," *Phys. Rev.*, vol. 79, no. 4, p. 046707, 2009.

[40] G. Wild and A. Traulsen, "The different limits of weak selection and the evolutionary dynamics of finite populations," *J. Theor. Biol.*, vol. 247, no. 2, pp. 382–390, 2007.

[41] S. Morita, "Extended pair approximation of evolutionary game on complex networks," *Progr. Theoret. Phys.*, vol. S Morita, no. 1, pp. 29–38, 2008.

[42] M. Newman, "Ego-centered networks and the ripple effect," *Social Netw.*, vol. 25, no. 1, pp. 83–95, 2003.

[43] P. Erdős and A. Rényi, "On random graphs I," *Publ. Math. Debrecen*, pp. 290–297, 1959.

[44] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[45] J. McAuley and J. Leskovec, "Learning to discover social circles in ego networks," in *Proc. Neural Inf. Process. Syst. Conf. (NIPS)*, 2012, pp. 1–9.

[46] J. Leskovec, Stanford Large Network Dataset Collection, [Online]. Available: http://snap.stanford.edu/data

[47] J. Leskovec, L. Backstrom, and J. Kleinberg, "Meme-tracking and the dynamics of the news cycle," in *Proc. 15th ACM SIGKDD Int. Conf. Knowl. Discov. and Data Mining (KDD)*, 2009, pp. 497–506.
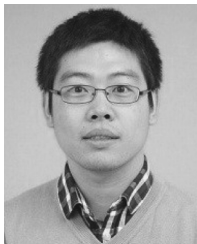
Chunxiao Jiang (S'09–M'13) received his B.S. degree in information engineering from Beijing University of Aeronautics and Astronautics (Beihang University) in 2008 and the Ph.D. degree from Tsinghua University (THU), Beijing in 2013, both with the highest honors. During 2011–2013, he visited the Signals and Information Group (SIG) at the Department of Electrical & Computer Engineering (ECE) of the University of Maryland (UMD) with Prof. K. J. Ray Liu. Dr. Jiang is currently a post-doctor in EE department of THU with Prof. Yong Ren. His research interests include the applications of game theory and queuing theory in wireless communication and networking and social networks.

Dr. Jiang received the Best Paper Award from IEEE GLOBECOM in 2013, the Beijing Distinguished Graduated Student Award, Chinese National Fellowship and Tsinghua Outstanding Distinguished Doctoral Dissertation in 2013.

Yan Chen (S'06–M'11) received the Bachelors degree from the University of Science and Technology of China in 2004, the M. Phil degree from Hong Kong University of Science and Technology (HKUST) in 2007, and the Ph.D. degree from University of Maryland College Park in 2011. His current research interests are in data science, network science, game theory, social learning and networking, as well as signal processing and wireless communications.

Dr. Chen is the recipient of multiple honors and awards including the best paper award from IEEE GLOBECOM in 2013, Future Faculty Fellowship and Distinguished Dissertation Fellowship Honorable Mention from the Department of Electrical and Computer Engineering in 2010 and 2011, respectively, Finalist of the Deans Doctoral Research Award from A. James Clark School of Engineering at the University of Maryland in 2011, and the Chinese Government Award for outstanding students abroad in 2011.

K. J. Ray Liu (F'03) was named a Distinguished Scholar-Teacher of the University of Maryland, College Park, in 2007, where he is Christine Kim Eminent Professor of Information Technology. He leads the Maryland Signals and Information Group conducting research encompassing broad areas of signal processing and communications with recent focus on cooperative and cognitive communications, social learning and network science, information forensics and security, and green information and communications technology.

Dr. Liu is the recipient of numerous honors and awards including the IEEE Signal Processing Society Technical Achievement Award and Distinguished Lecturer. He also received various teaching and research recognitions from the University of Maryland including university-level Invention of the Year Award; and Poole and Kent Senior Faculty Teaching Award, Outstanding Faculty Research Award, and Outstanding Faculty Service Award, all from the A. James Clark School of Engineering. An ISI Highly Cited Author, Dr. Liu is a Fellow of IEEE and AAAS.

Dr. Liu is Past President of the IEEE Signal Processing Society where he has served as Vice President Publications and Board of Governor. He was the Editor-in-Chief of the *IEEE Signal Processing Magazine* and the founding Editor-in-Chief of *EURASIP Journal on Advances in Signal Processing*.