Privacy Protection: A Community-Structured Evolutionary Game Approach

Jun Du*†, Chunxiao Jiang*, Shui Yu‡, Kwang-Cheng Chen§ and Yong Ren*
*Department of Electronic Engineering, Tsinghua University, Beijing, 100084, China
†Tsinghua National Laboratory for Information Science and Technology, Tsinghua University, Beijing 10008, China
‡School of Information Technology, Deakin University, Burwood, VIC 3125, Australia
§Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan
E-mail: {blgdujun, chx.jiang}@gmail.com, syu@deakin.edu.au, ckc@ntu.edu.tw, reny@tsinghua.edu.cn

***Abstract*—Users of social networks can be connected with each other by different communities according to professions, living locations and personal interests. As each user on the social network platforms stores and shows a large amount of personal data, the privacy protection raises as a major concern. This paper establishes a game theoretic framework to model users' interactions to influence users' strategies to take the privacy protection or not. To model the relationship of user communities, we introduce the community-structured evolutionary dynamics. Users' interactions can only happen among those who have at least one common community. Then we analyze the dynamics of users' privacy protection behavior based on the proposed community structured evolutionary game theoretic framework. Results show that social network managers need to provide appropriate security service b and payment mechanism c to ensure that cost performance b/c is larger than the critical cost performance, which can promote the spread of the privacy security behavior over the network. Moreover, results can help to design appropriate structure of the social network and control the convergence speed that all users take the privacy protection.**

I. Introduction

Recently, we have seen unprecedented development of the social network applications. Emerging online social networks, such as Facebook and Twitter, are inherently designed to enable people to distribute and share personal and public information [1]. In addition, social connections among friends, cooperators, and even strangers of peculiar interest are established via these online social networks. However, as allowing their users to host large amounts of personal data on their platforms, important concerns regarding the security, especially the privacy security, of the user related information raise with the rapid technological development in areas such as social networking, online applications and cloud commuting [2], [3]. How to protect users' personal information to ensure the privacy security, which can gain users' trust and encourage them to participate into the social network, has become one of critical problems for social network managers.

In response, many privacy protection mechanisms have been studied from many aspects to improve the security of users' data [4], [5], [6]. Game theory and evolutionary game theory (EGT) have been applied to comprehend and to interpret the interactions among users regarding personal information security. In [7], researchers formulated a non-cooperative security information sharing game. A model of evolutionary game between social network sites and users was established from the perspective of privacy concerns in [8]. In [9], the network security and privacy in multi-hop network were modeled as a game among the nodes. On the other hand, the influence of the users' behaviors also plays an important role on the selection and spreading of privacy protection over social networks. Results of [10] demonstrated that the messages directly influenced political self-expression, information seeking and real world voting behavior of millions of people. It was suggested in [11] that if the goal of policy is to adequately protect privacy, then we need to protect individuals with minimal requirement of informed and rational decision making that includes a baseline framework of protection.

Most current research on privacy protection and behavior spreading considers a social network of a regular, random, and flattened topology. Then users connect with each other, and the influence of users' behaviors is spread over the entire social network accordingly. However, in a real social network, the relationship among users are much more complicated than above models. Moreover, the interaction and influence between any two users largely depends on how close the relationship between these two users is. In this research, we model the population of social networks as a community structure to discuss the connections of users in a more appropriate and accurate way. By this model, we may successfully analyze the spreading of the privacy protection behavior over the network.

The rest of this paper is organized as follows. In Section II, the community structure based evolutionary game formulation is described. The privacy protection among users belonging to K communities is analyzed in Section III. Simulations are shown in Section IV, and conclusions are drawn in Section V.

II. Community Structured Evolutionary Game

A. Basic Conception of Evolutionary Game

Consider an evolutionary game (EG) with r strategies $\chi = \{1, 2, \cdots, r\}$ and a payoff matrix $\mathbf{U}$, which is an $r \times r$ matrix with entry u_{mn} denoting the payoff for strategy m versus strategy n. The system state of the game can be denoted as $\mathbf{p} = [p_1, p_2, \cdots, p_r]^{\mathrm{T}}$. In this case, the average mean payoff within a population in state $\mathbf{q} = [q_1, q_2, \cdots, q_r]'$ against a population in state $\mathbf{p}$ is $\mathbf{q}'\mathbf{U}\mathbf{p}$. Then a state $\mathbf{p}^*$ is an *Evolutionary Stable*

State (ESS), if and only if $\mathbf{p}^*$ satisfies following conditions for all different states $\mathbf{q} \neq \mathbf{p}$ [12]: $\mathbf{q}'\mathbf{U}\mathbf{p}^* \leq \mathbf{p}^{*\prime}\mathbf{U}\mathbf{p}^*$, and if $\mathbf{q}'\mathbf{U}\mathbf{p}^* = \mathbf{p}^{*\prime}\mathbf{U}\mathbf{p}^*$, $\mathbf{p}^{*\prime}\mathbf{U}\mathbf{q} > \mathbf{q}'\mathbf{U}\mathbf{q}$. An approach to find the ESS is to find the stable point $\mathbf{p}^* = \arg_{\mathbf{p}}(d\mathbf{p}/dt = 0)$ of the network state dynamics, if it exists [13].

The users with evolutionary actions are considered as discrete and non-overlapping updated generations, and the amount of users is a constant. All users update at the same time. Users imitate other users' communities and security behaviors in the new update proportional to their fitness [13], which means that if one user has a higher fitness, other users tend to imitate his current strategy and community memberships in the following update with a high probability. Define that the inheritance of the community memberships occurs with deviate rate v, and the inheritance of the strategies subjects to deviate rate u. A user adopts the imitated user's community memberships with probability $1-v$, or adopts a random configuration, which includes that of the imitated user, with probability v. Similarly, an offspring adopts the imitated user's strategy with probability $1-u$, or adopts a random strategy with probability u.

B. Community-Structured Evolutionary Game Formulation

Assume that a social network can provide a higher grade of security, i.e., privacy protection, for users' privacy. This additional security service can be obtained only by those users who have accepted terms ruled by the network, such as that users have to complete real-name authentication or pay for the service. We consider that this service is not mandatory.

In current social networks, users are allowed to join multiple but a limited number of communities according to their professions or personal interests. For instance, the Facebook Groups establish different communities. Therefore, each user hold multiple friendship relations with the users in the same community. The closeness of the relationship between two uses can be measured by the number of the communities they sharing. Assume that users are allowed to change communities, Consider that user interactions can only happen between individuals belonging to the same community. In addition, some of users' information is accessed only to their friends. Based on these premises above, we assume that the information of both of the user and his friends can be protected by the network to some extent if the user take the privacy protection, even if his friends do not take the same action. Meanwhile, the information of his friends not taking the privacy protection can also be protected by the network although without any contributions of these users. Taking Facebook for instance, the Timeline of an user can only be seen by his friends. If the user's friend take the privacy protection or some other information security services, the accessible information of this user can be also protected on some level.

The privacy protection over a social network shares fundamental similarities with the strategy updating in the community-structured EGT. Social network users are the players in the EG. Each of these users has two possible strategies, i.e., to take or not take the privacy protection provided by the network, denoted by $\mathbf{S}_p$ and $\mathbf{S}_n$, respectively. The strategy taking the privacy protection can be considered as the security behavior, and otherwise, insecurity behavior. Meanwhile, a users' payoff matrix can be defined as

$$\begin{array}{cc} & \begin{array}{cc} \mathbf{S}_p & \mathbf{S}_n \end{array} \\ \begin{array}{c} \mathbf{S}_p \\ \mathbf{S}_n \end{array} & \begin{pmatrix} \beta b - c & b - c \\ b & 0 \end{pmatrix} \end{array}, \tag{1}$$

where $b > 0$ is the baseline security benefit received by the user resulting from that he/she or his friend takes the privacy protection. $c > 0$ denotes the cost that users taking the privacy protection need to pay for the protection service. In addition, when both of the interacted users take strategy $\mathbf{S}_p$, two of them will obtain higher safety benefit βb, where $\beta > 1$. The payoff will be zero when neither of the interacted friends selects the privacy protection service, then no pay or gain for them.

Based on the definitions above, ratio b/c or $\beta b/c$, which can be defined as the *cost performance*, is a crucial parameter. It can help social network managers to make appropriate security service level and payment mechanism to encourage users to accept the security service, and then promote the spreading of the security behavior. So it is necessary to investigate the interaction between users with opposite security behavior, and figure out the question that whether dynamics on a community structured population allows the evolution of security behavior.

Consider a social network with N users distributed over M communities. Each user belongs to K ($K \leq M$) communities. In addition, each user has a strategy index $s_i \in \{0, 1\}$, which is defined as that $s_i = 1$ when user i takes the privacy protection strategy $\mathbf{S}_p$, or $s_i = 0$, otherwise. Then the social network state can be given by a strategy vector $\mathbf{s} = [s_1, s_2, \cdots, s_N]$ and a matrix $\Theta_{N\times M}$, whose entry θ_{im} ($i = 1, 2, \cdots, N$, $m = 1, 2, \cdots, M$) is 1 if user i belongs to community m, and $\theta_{im} = 0$, otherwise. Define $\Theta = [\theta_1, \theta_2, \cdots, \theta_N]^{\mathrm{T}}$, where θ_i is the vector giving the community membership of user i. Then the number of communities that users i and j having in common can be expressed by $\theta_i \cdot \theta_j$, where $\cdot$ denotes the dot product.

We assume that two users interact as many times as the number of their common communities, and the self-interaction is ignored in our work. For each interaction, the fitness of user i is determined by (1). Then the total fitness of user i of the community-structured social network can be written as

$$\pi_i = 1 + \alpha \sum_{j \neq i} (\theta_i \cdot \theta_j) \left[(\beta - 2) b s_i s_j + (b - c) s_i + b s_j \right], \tag{2}$$

where α represents the relative contribution of the game to fitness. The cases $\alpha \to 0$ and $\alpha = 1$ denote the weak selection and strong selection, respectively. An example of the update process is shown in Fig. 1, in which user U_1 picks user U_2, and adopts U_2's security strategy and community associations.

III. Privacy Protection among Users Belonging to K Communities

A. Evolution of Security Behavior on Communities

Assume that users can change their community memberships and security behaviors to get better user security experi-

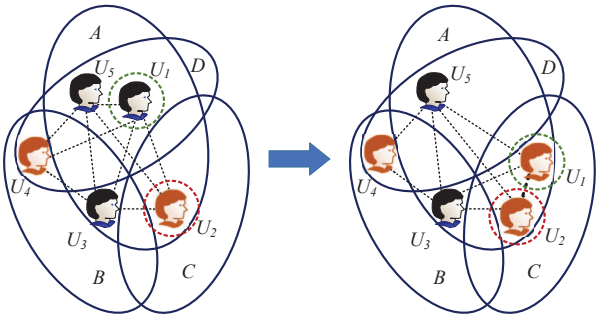

Fig. 1. An example of security strategy and associations evolution over a social network with a community-structured population. $N = 5$, $M = 4$ and $K = 2$. The broken lines indicate the weighted interaction.

ence. This change can be considered as the update. By times of updates, the behavior with higher fitness tends to be imitated and can spread among the users in the network. Moreover, the community and strategy deviation can also make sense on the behavior spreading. On the one hand, when an user imitates anther one for a better security experience, he/she only imitates the community membership. On the other hand, a user might only imitate another user's security behavior but not change some or all of his previous communities because of interests. These two situations bring security behavior (strategy) deviation and community deviation, respectively, as mentioned in Section II-A. We still use u and v to denote the rates of strategy deviation and community deviation, respectively.

Define the user state as $(p, 1-p)$, where p is the frequency of the users selecting strategy $\mathbf{S}_p$. Our ultimate goal is to derive the ESS $(p^*, 1-p^*)$ that ensures the evolution of security behavior, i.e., users select strategy $\mathbf{S}_p$ more frequently than $\mathbf{S}_n$. To find out the ESS of the system state dynamic, we need to analyze the effect of imitation and deviation on the average change in p. Since that the average value of p is a constant, the two effects must cancel [14]. Then we can get

$$\langle \hat{p} \rangle_{\text{imi}} + \langle \hat{p} \rangle_{\text{dev}} = 0, \tag{3}$$

where $\langle \hat{p} \rangle_{\text{imi}}$ and $\langle \hat{p} \rangle_{\text{dev}}$ denote the effect of imitation and deviation, respectively, and they are both the continuous functions of α. Consider the weak selection situation that $\alpha = 0$, and

$$\langle \hat{p} \rangle_{\text{imi}} = 0 + \alpha \langle \hat{p} \rangle_{\text{dev}}^{(1)} + o\left(\alpha^2\right), \tag{4}$$

where $\langle \hat{p} \rangle_{\text{imi}}^{(1)}$ is the first derivative of $\langle \hat{p} \rangle_{\text{imi}}$ with $\alpha = 0$, and the third equality is according to Taylor's Theorem. When $\langle \hat{p} \rangle_{\text{imi}}^{(1)} > 0$, the amount of users who take the privacy protection due to the imitation increases, which means that the user's decision tends to be the security behavior. Contrarily, if $\langle \hat{p} \rangle_{\text{imi}}^{(1)} < 0$, the user's decision tends to be not taking the privacy protection.

B. Finding the Critical Ratio

In order to obtain the critical parameter value of cost performance, we must have $\langle \hat{p} \rangle_{\text{imi}}^{(1)} = 0$. In this work, we analyze the neutral stationary state and get the general expression of cost performance. We provide the critical cost performance $(\beta b/c)^*$ in the limit of weak selection as Theorem 1.

Theorem 1. *In a social network with N users, every user belongs to exactly K communities. There are two strategies $\mathbf{S}_p$ and $\mathbf{S}_n$ for users. Interactions are only allowed among users sharing communities in common. For each user, the payoff matrix is given by (1). The deviate rates of community membership imitation and strategy imitation are given by v and u, respectively. The critical cost performance that keeps the neutral stationary state is given by*

$$\left(\frac{\beta b}{c}\right)^* = 1 + \frac{\mu+v+3}{\mu+v+1} \cdot \frac{Kv\left(\mu+v+2\right)+M\left(\mu+1\right)}{Kv\left(\mu+v+2\right)+M\left(\mu+2v+3\right)}, \tag{5}$$

where $v = 2Nv$ and $\mu = 2Nu$. For $\mu \to 0$, we have

$$\left(\frac{\beta b}{c}\right)^* = 1 + \frac{v+3}{v+1} \cdot \frac{Kv\left(v+2\right)+M}{Kv\left(v+2\right)+M\left(2v+3\right)}. \tag{6}$$

For the space limitation, we provide the detailed proof in [15]. Theorem 1 gives the critical cost performance $(b/c)^*$ or $(\beta b/c)^*$. In the equilibrium distribution of the imitation-deviation process, if the cost performance exceeds this critical value, the users in the social network will select the strategy of privacy protection more frequently than the other strategy, i.e., taking no privacy protection, which will promote the diffusion of security behaviors among the network. Moreover, consider $(\beta b/c)^*$ provided in Theorem 1 as a function of K/M, and we take the derivative of $(\beta b/c)^*$ with respect to K/M, then we get $\frac{\partial(\beta b/c)^*}{\partial(K/M)} > 0$. So $(\beta b/c)^*$ increases with increasing K/M. Hence, for a social network with M communities, the best choice for social network managers to set the minimize $(\beta b/c)^*$ is allowing their users to belong to only one community, i.e., $K = 1$.

C. Privacy Protection with L-Triggering Game

In a social network, the interaction between two users sometimes depends on the strength of their connection, which could be measured by the number of communities that they have in common. Specifically, user i and j are sharing a close relationship, which means that they have many interested communities in common. As a result, most information of user j is accessible for user i. In this case, if user i selects the privacy protection, user j's personal information even privacy information can be protected to a great extent. Conversely, if the amount of the two users' common communities is really small, for instance, user i and j coming from different countries just join the same travel community because of their annual leaves, then the relationship between the two users is actually quite weak and there is little personal information can be accessed for each other. In this case, user j cannot benefit from user i's selection of privacy protection.

In response, we generalize the model, in which the users' interaction happens as long as they have at least one community in common, into a *L-triggering game* situation in this section. In the extended model, users only influence each other if they have at least a minimum number of common communities, L. In a social network, if an user taking the privacy protection i meets another user j in $\theta_i \cdot \theta_j$ communities, then i interacts $\theta_i \cdot \theta_j$ times if $\theta_i \cdot \theta_j \geq L$, otherwise, the game between them is

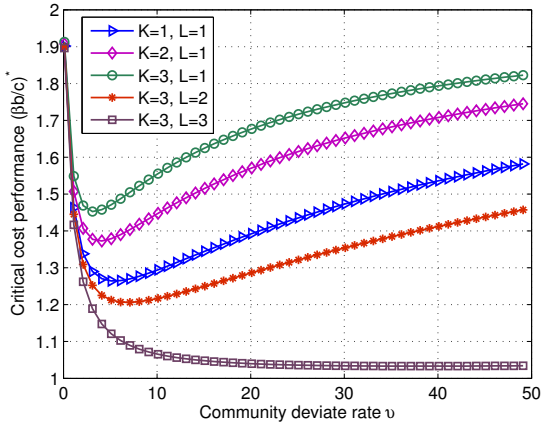

Fig. 2. Critical cost performance $(\beta b/c)^*$ versus community deviate rate υ.

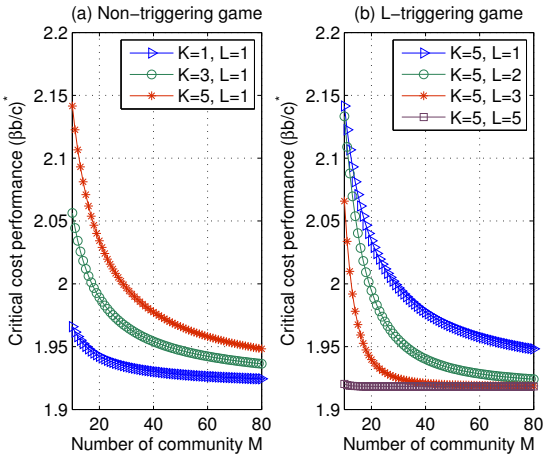

Fig. 3. Critical cost performance $(\beta b/c)^*$ versus communities number M.

not triggered. We call this mechanism as *L-triggering game*. We notice that $L=1$ degenerates to the previous model. The analysis of cost performance at the end of this section indicates that large values of L lead to that users with security behavior are more imitative in choosing with whom to imitate. Next, we will analyze the impact of L-triggering game on the critical cost performance.

Given $1 \leq L \leq K$. When $L=1$, the model is same as of Section III (B). Then equation (2) can be rewritten as

$$\pi_i = 1+\alpha \sum\nolimits_{j \neq i} \chi_{ij}\left(\theta_i \cdot \theta_j\right)\left[(\beta-2)\, b s_i s_j + (b-c)\, s_i + b s_j\right], \quad (7)$$

where $\chi_{ij}=1$ if $\theta_i \cdot \theta_j \geq L$, and $\chi_{ij}=0$, otherwise.

According to some mathematical derivations, we obtain the critical cost performance formulated in equation (6) as

$$\left(\frac{\beta b}{c}\right)^* = 1+\frac{\upsilon+3}{\upsilon+1}\cdot\frac{\hat{K}\upsilon\left(\upsilon+2\right)+M}{\hat{K}\upsilon\left(\upsilon+2\right)+M\left(2\upsilon+3\right)}, \quad (8)$$

in case that $N \to \infty$ and $\mu \to 0$. In (8),

$$\hat{K} = \frac{M}{K}\sum\nolimits_{i=L}^{K} i\binom{K}{i}\binom{M-K}{K-i}/\binom{M}{K}. \quad (9)$$

Due to the space limitation, we also provide the derivation in [15]. We can notice that $\hat{K}=K$, if $L=1$. Comparing with equation (6), the expressions of $(\beta b/c)^*_{\min}$ for non-triggering game and L-triggering game are much the same, except that $\hat{K} \leq K$, and the equality holds up if and only if $L=1$.

IV. Simulation Results

In this part, we perform numerical simulation experiments to analyze properties and performances of the critical cost performance and its influential factors such as the community deviate rate and number of communities of the social network.

First, the community deviate rate υ reflects the subjective selectivity for community memberships. If users select communities depending on their own interest mostly, but not on those users with high fitness, then υ is large. Then we analyze the effect of the community deviate rate $\upsilon=2N\upsilon$ for different K and L. The population of the social network is large, i.e., $N=10^4$ $(N \to \infty)$, and the number of communities is $M=20$. Set the strategy deviate rate as $u=10^{-4}$ $(u \to 0)$. Consider the population of the network is constant. Simulation results of non-triggering and L-triggering game are shown in Fig. 2. As shown in the results, the critical cost performance $(\beta b/c)^*$ is a U-shaped function of community deviate rate υ. When υ is small, $(\beta b/c)^*$ tends to be large and all users belong to the same community. Conversely, when υ is large, the community affiliations cannot persist for a long time.

As shown in Fig. 2, we notice that for a fixed M, small values of K can facilitate the evolution of the security behavior, which means that the selection of taking the privacy protection is promoted in the evolution process. Consequently, when M is given, the best choice for users is to belong to $K=1$ community. With the increasing of K, it is hard for users taking the privacy protection to avoid the exploitation by users not taking the privacy protection. But according the results of the L-triggering game situation, for $K=3$, if $L=2$ or $L=3$, the critical cost performance is smaller than $K=1$. These results indicate that belonging to more communities, i.e., $K>1$, can also facilitate the evolution of the security behavior when the game only happen if users have a certain minimum number of common communities L.

As shown in Fig. 3, the cost performance decreases as the number of communities M increases. These results indicate that more communities are helpful for the spreading of security behavior, which means that adding community number will help users to take the privacy protection more frequently.

V. Conclusion

In this paper, we propose a community-structured EGT framework to model and analyze the privacy protection behaviors of social network users. We obtain the critical cost performance, which is an important parameter that can help to design incentive mechanisms to facilitate the privacy protection behavior among users. Simulation results demonstrate that the proposed theoretic framework is effective in modeling the users' relationship and security behavior.

Acknowledgment

This research was supported by NSFC China under projects 61371079, 61271267 and 91338203.

REFERENCES

[1] E. Serrano, C. A. Iglesias, and M. Garijo, "A survey of twitter rumor spreading simulations," in *Computational Collective Intelligence*. Springer, 2015, pp. 113–122.

[2] I. Krontiris, M. Langheinrich, and K. Shilton, "Trust and privacy in mobile experience sharing: future challenges and avenues for research," *Communications Magazine, IEEE*, vol. 52, no. 8, pp. 50–55, Aug. 2014.

[3] C. Bettini and D. Riboni, "Privacy protection in pervasive systems: State of the art and technical challenges," *Pervasive and Mobile Computing*, vol. 17, pp. 159–174, Feb. 2015.

[4] J. Adebayo and L. Kagal, "A privacy protection procedure for large scale individual level data," in *Intelligence and Security Informatics (ISI), 2015 IEEE International Conference on*. Baltimore, MD, May 2015, pp. 120–125.

[5] R.-H. Hwang, Y.-L. Hsueh, and H.-W. Chung, "A novel time-obfuscated algorithm for trajectory privacy protection," *Services Computing, IEEE Transactions on*, vol. 7, no. 2, pp. 126–139, Jun. 2014.

[6] C. L. Apicella, F. W. Marlowe, J. H. Fowler, and N. A. Christakis, "Social networks and cooperation in hunter-gatherers," *Nature*, vol. 481, no. 7382, pp. 497–501, 2012.

[7] D. Tosh, S. Sengupta, C. Kamhoua, K. Kwiat, and A. Martin, "An evolutionary game-theoretic framework for cyber-threat information sharing," in *Communications (ICC), 2015 IEEE International Conference on*. London, UK, Jun. 2015, pp. 7341–7346.

[8] W. Lian-ren and C. Xia, "Modeling of evolutionary game between sns and user: From the perspective of privacy concerns," in *Management Science & Engineering (ICMSE), 2014 International Conference on*. Arunachal Pradesh, India, May 2014, pp. 115–119.

[9] C. Kamhoua, N. Pissinou, K. Makki *et al.*, "Game theoretic modeling and evolution of trust in autonomous multi-hop networks: Application to network security and privacy," in *Communications (ICC), 2011 IEEE International Conference on*. Kyoto, Japan, Jun. 2011, pp. 1–6.

[10] R. M. Bond, C. J. Fariss, J. J. Jones, A. D. Kramer, C. Marlow, J. E. Settle, and J. H. Fowler, "A 61-million-person experiment in social influence and political mobilization," *Nature*, vol. 489, no. 7415, pp. 295–298, Sept. 2012.

[11] A. Acquisti, L. Brandimarte, and G. Loewenstein, "Privacy and human behavior in the age of information," *Science*, vol. 347, no. 6221, pp. 509–514, Jan. 2015.

[12] C. Jiang, Y. Chen, and K. R. Liu, "Graphical evolutionary game for information diffusion over social networks," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 8, no. 4, pp. 524–536, Aug. 2014.

[13] R. Cressman, *Evolutionary dynamics and extensive form games*. MIT Press, 2003, vol. 5.

[14] C. E. Tarnita, T. Antal, H. Ohtsuki, and M. A. Nowak, "Evolutionary dynamics in set structured populations," *Proceedings of the National Academy of Sciences*, vol. 106, no. 21, pp. 8601–8604, May 2009.

[15] J. Du, C. Jiang, S. Yu, K.-C. Chen, and Y. Ren, "Appendix for privacy protection: A community-structured evolutionary game approach," [Online]: http://www.jiangchunxiao.net/egsappendix.pdf.