A review on graphical evolutionary game for information diffusion on social networks

Benliu Qiu and Ningxuan Zhang
School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan Province, China

Abstract

Purpose – With the recent development of science and technology, research on information diffusion has become increasingly important.

Design/methodology/approach – To analyze the process of information diffusion, researchers have proposed a framework with graphical evolutionary game theory (EGT) according to the theory of biological evolution.

Findings – Through this method, one can study and even predict information diffusion.

Originality/value – This paper summarizes three existing works using graphical EGT to discuss how to obtain the static state and the dynamics of information diffusion in social network.

Keywords Crowd-sourced design and engineering, Information diffusion, Crowdsourced Big Data and analytics, Evolutionary game, Models and methods for Crowd Science and Engineering

Paper type General review

1. Introduction

Over the past several decades, accompanied by the rapid development of the internet and mobile communications, social networks are becoming increasingly popular among numerous individuals. Every second, millions of users are interacting, which is giving rise to the development of various kinds of information. It is therefore critical to understand how such a huge amount of information diffuses over social networks for the development of social networks.

Many researchers have been studying information diffusion from different perspectives. Originally, information diffusion was proposed to solve the problem of attacks of computer virus (Pastor-Satorras and Vespignani, 2001) (Gruhl et al., 2004). People later realized that information diffusion could be an independent research topic. Early works about information diffusion focused on the dynamics of information diffusion from both the
macroscopic and microscopic perspectives (Gruhl et al., 2004). Later, researchers studied how users’ behaviors were affected by social networks (Centola, 2010). Machine learning was used to find the global influence of individuals in diffusion (Yang and Leskovec, 2010). How to restrain private information diffusion by important information links was investigated by Kimura et al. (2009) and Ilyas et al. (2011). The authors discussed how to find the nodes that have more influence over others (Kimura et al., 2007). Recently, as Facebook and Twitter have become increasingly popular, an analysis was conducted to predict future trends on the basis of data obtained from Twitter (Yang and Counts, 2010) (Leskovec). Game theory has also been used to analyze this process (Chen and Liu, 2012). The graphical evolutionary game theory (EGT) framework was proposed to model and analyze interaction among users in social network (Jiang et al., 2014a, 2014b), and the coordination game was utilized to analyze the diffusion process (Vega-Redondo, 2007). The stability of information diffusion was obtained by means of different methods (Wang et al., 2010a, 2010b, 2010c; Kim and Yoneki, 2012). It was shown that users with different number of followers can have different influences (Beevolve, 2012).

Information diffusion over social networks is similar to biological evolution, and thus, the graphical EGT can be used to analyze and model how information diffuses over social networks. EGT has been widely adopted in various areas including image interpolation (Chen et al., 2011), as well as wireless communications and networking (Wang et al., 2010a, 2010b, 2010c; Jiang et al., 2013; Tembine et al., 2010; Wieck et al., 2011). In this paper, we summarized the key results of studies (Jiang et al., 2014a, 2014b) and (Cao et al., 2016) to illustrate how the graphical EGT can be used to model information diffusion.

The rest of this paper is organized as follows. In Section 2, information diffusion is discussed in detail. Some basic concepts about graphical EGT are given in Section 3. Then in Section 4, the model built with graphical EGT is discussed. Static and dynamic diffusion processes with the corresponding experimental results are shown in Section 5. Section 6 draws the conclusion of this paper.

2. Information diffusion

To understand information diffusion, the information should be defined. When weather prediction tells us it will be sunny tomorrow or teachers solve our confusion about an incomprehensible concept, we instinctively feel that we have received information. It is widely agreed that information is any entity or form that resolves uncertainty or provides an answer for any question. Despite information being present since the birth of humankind, the measurement of information was discussed only in in 1948 by C. E. Shannon in the book *A Mathematical Theory of Communication*.

Just as the information has been with human for so long, so is the information diffusion. This is because where there is information, there is information diffusion. With the rapid development of the internet and mobile communications, information diffusion over social networks is faster and broader. For example, sham information has great possibilities of becoming a rumor as soon as it is released in a social network. In a traditional social network, there may be adequate time for response and stopping the rumor. However, in contemporary social network, more efforts are needed to stop rumors. Studying information diffusion over social network can help us discover the common characteristics of communities for precise commercial advertisement and pointed political canvassing. In all, the study of information diffusion can be applied to various domains. There are two examples. Donald Trump, the President of the USA, has the habit of tweeting for the country. As the president of the world’s superpower, his words on the internet have the ability of creating waves. If we can predict how the information released by Trump is
diffused on the internet or among the people, then we can better evaluate the influence caused of his tweets. What does it mean? It means, in the financial world, we will macroscopically be able to more quickly and accurately predict fluctuations in the stock, which can prove to be financially beneficial. Except predicting the diffusion states, we have the ability to deduce the characteristic of a group of people who are forwarding or not through known diffusion states at any moment. This can be used to tell candidates to give corresponding speeches to different groups of voters. It is also possible that the popularity rating can be raised. Besides this, the model of users, which is obtained by analyzing the existing diffusion actions of users in the past, can be applied to advertise on the internet more accurately, which may create billions of wealth, as it has done for Google.

3. Graphical evolutionary game theory
The basic concepts of graphical EGT will be introduced in this section (Smith, 1982), which will be used to analyze information diffusion in the next section (Jiang et al., 2014a, 2014b) and (Cao et al., 2016).

In EGT, a game is played for many times by players who are chosen randomly from a large population. In this theory, a player changes his strategy according to others’ strategies and his characteristics. We call the dynamics as “replicator dynamics” which is the most important concept in EGT. Replicators, who can represent the population, may reproduce his strategy under some specific rules. For example, there is new information about a football star. One person who is a football fan may originally prefer to forward this information, but when he finds that all of his friends are not concerned about such information, he may change his strategy from forwarding to not forwarding. In this example, we can see that the strategies of neighbors can influence one’s decision. Different from traditional game theory, EGT pays more attention to the dynamics than to stability of the strategies of the whole population. In EGT, all players will have some periods of strategic interactions to reach a final equilibrium, known as the evolutionary stable state (ESS).

In the social network, people may not interact with all users. Therefore, a graph is generally used to represent the relationship among users. In this graph, we use “fitness” (Nowak and Sigmund, 2004) to describe players’ interactions with others, which is relative to their payoff. Players change their strategies according to three strategy update rules (Ohtsukia and Nowak, 2006), including birth–death (BD), death–birth (DB) and imitation (IM) rules. Figure 1 illustrates the general process of these three update rules (Jiang et al., 2014a, 2014b). In Figure 1(a), we can see in the BD update rule, a node is chosen from the population with a probability according to their fitness. Then the strategy of the chosen node will replace one of its neighbor’s strategy. In the DB update rule, we randomly choose a node to abandon its strategy and get one of its neighbor’s strategy with the probability according to their fitness, which is shown in Figure 1(b). In IM update rule, every node chooses one action: either it abandons or retains its strategy with the probability according to their fitness, as shown in Figure 1(c).

4. EGT formulation of information diffusion
Let us consider how to translate a typical social network with the language of graphical EGT. Any users in the social network can be regarded as a node in a graph. The relationship between users corresponds to the edges of graph. In a word, a graph structure can be used to describe the social network topology. Because information diffusion is studied over the social network, it is assumed that any user has only two strategies to choose from: to forwarding information or to not (S_f or S_n). Here, a question can be raised naturally: How do we describe the utility from forwarding or not? As has been done in graphical EGT, fitness
Nowak and Sigmund, 2004) can be used to represent the utility from forwarding or not, which is defined as \( \pi = (1 - \alpha)B + \alpha U \), where \( U \) is the payoff matrix and \( B \) is the user’s baseline fitness. In the study of information diffusion, the payoff matrix can be written as:

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

where \( u_{ff} \) denotes the payoff when one user and his neighbor both use strategy \( S_f \); \( u_{fn}, u_{nf}, \) and \( u_{nn} \) are defined in a similar way. Note that the payoff matrix is considered symmetric; therefore, the central user and his neighbor both have the same fitness when they use the reverse strategy, i.e. \( u_{ff} = u_{nf} \). The payoffs are further normalized within interval \((0, 1)\), which means \( 0 < u_{ff}, u_{fn}, u_{nn} < 1 \). The payoff can be regarded as the popularity of a user or the hit rate of a website by forwarding this information. For example, if the information is about recent hot news, it is known that forwarding it will get more popularity than not forwarding. So under this circumstance, we have \( u_{ff} > u_{fn} > u_{nn} \). On the contrary, if the information is useless for most of us, \( u_{ff} < u_{fn} < u_{nn} \) can be obtained.

The decision of users is determined by not only the payoff but also by the user’s internal property, i.e. the baseline fitness \( B \). The parameter \( \alpha \) in the fitness is a representation of the selection intensity. It is within interval \((0, 1)\). When it is close to 0, the corresponding selection is a weak selection (Ohtsuki et al., 2007), while when \( \alpha \rightarrow 1 \), it denotes a strong selection, in which the payoff determines the fitness. The results shown here are based on the assumption of the weak selection. The reason is that every user is involved in so many games that any game just influences a user a little.
Given a social network, let \( p_f \) and \( p_n \) denote the percentages of users to forwarding or not, respectively, in the population. Similarly, let \( p_{ff}, p_{fn}, p_{nf}, \) and \( p_{nn} \) denote the percentages of edges, where both users are forwarding, one is forwarding and the other is not and both users are not forwarding, respectively. With the basic knowledge of the probability theory, one has:
\[
\begin{align*}
&\ p_n + p_f = 1, \quad p_{ff} + p_{fn} + p_{nf} = 1, \quad p_f p_{nf} = p_{nf} p_f, \quad p_{ff} = p_f p_{fn}, \\
&\ p_{nn} = p_n p_{nf}, \quad p_{fn} = p_f p_{nf} + p_n p_{fn}.
\end{align*}
\]
It is obvious that all variables above can be denoted by \( p_f \) and \( p_{ff} \).

5. Theorems for information diffusion over social networks
To characterize the information diffusion over social network with graphical EGT, the two most important things are the evolutionary stable states (ESSs) and the evolutionary dynamics. As discussed above, the states of a social network can be determined by \( p_f \) and \( p_{ff} \). Thus, the following analysis will focus on \( p_f, p_{ff}, \dot{p}_f \) and \( \dot{p}_{ff} \). The first result is about the population dynamics \citep{jiang2014a, jiang2014b}.

**Theorem 1**: The population dynamics of information diffusion over a uniform degree network under BD strategy update rule and weak selection scenario can be described as follows:
\[
\dot{p}_f = \alpha^2 (k - 2) \frac{k - 1}{(k - 1)^2} p_f (1 - p_f) (a_1 p_f + b_1) + o(\alpha^2),
\]
where
\[
\begin{align*}
& \quad a_1 = (k - 2)(u_{ff} - 2u_{fn} + u_{nn}), \\
& \quad b_1 = u_{ff} + (k - 2)u_{fn} - (k - 1)u_{nn}.
\end{align*}
\]

**Theorem 2**: The population dynamics of information diffusion over non-uniform degree networks under the BD strategy update rule and weak selection scenario can be described as follows:
\[
\dot{p}_f = \alpha^2 (k - 1)(\frac{k^2 - 2 \bar{k}}{k^2 - \bar{k}})^2 p_f (1 - p_f) (a_2 p_f + b_2) + o(\alpha^2),
\]
where
\[
\begin{align*}
& \quad a_2 = (\bar{k}^2 - 2 \bar{k})(u_{ff} - 2u_{fn} + u_{nn}), \\
& \quad b_2 = \bar{k} u_{ff} + (\bar{k}^2 - 2 \bar{k})u_{fn} - (\bar{k}^2 - \bar{k})u_{nn}.
\end{align*}
\]
From equations (1) and (2), it can be seen that \( \dot{p}_f \) is not related with \( N \), i.e. the network has the scale-free property, which indicates that the results do not depend on the number of users in the whole network as long as it is large enough. It can be seen that the population dynamics in the uniform and non-uniform degree networks are similar. The only difference between them is \( k \) and \( k^2 \) are changed to \( \bar{k} \) and \( \bar{k}^2 \) in the non-uniform network, which means that the non-uniform network can be treated as a uniform network. Although the above two theorems are based on the BD update rule, it has been shown that the population dynamics of information diffusion over uniform degree networks under the BD strategy update rule, the DB strategy update rule and the imitation strategy update rule are equivalent when the
network degree is sufficiently large and under the weak selection scenario (Jiang et al., 2014a, 2014b).

To validate the theorems over a homogeneous social network, the following payoff matrices PM1, PM2, PM3 and PM4 are used to derive the theoretical results and compare with simulation results.

Case 1:
PM1: $u_{ff} > u_{fn} > u_{nn}$

\[
\begin{array}{cc}
S_f & S_n \\
S_f & 0.8 & 0.6 \\
S_n & 0.6 & 0.4 \\
\end{array}
\]

Case 2:
PM2: $u_{fn} > u_{ff} > u_{nn}$

\[
\begin{array}{cc}
S_f & S_n \\
S_f & 0.6 & 0.8 \\
S_n & 0.8 & 0.4 \\
\end{array}
\]

Case 3:
PM3: $u_{fn} > u_{nn} > u_{ff}$

\[
\begin{array}{cc}
S_f & S_n \\
S_f & 0.4 & 0.8 \\
S_n & 0.8 & 0.6 \\
\end{array}
\]

Case 4:
PM4: $u_{nn} > u_{fn} > u_{ff}$

\[
\begin{array}{cc}
S_f & S_n \\
S_f & 0.4 & 0.6 \\
S_n & 0.6 & 0.8 \\
\end{array}
\]

For Case 1: $u_{ff} > u_{fn} > u_{nn}$, the user and his neighbor who both forward information can achieve the largest payoff, so it can be predicted that the users will choose the forwarding strategy, making all users in the network forward it gradually. For Case 4, $u_{nn} > u_{fn} > u_{ff}$, the final state of the network will be the complete opposite to that of Case 1, i.e. nobody will choose to forward information. The final stable states of Case 2 and Case 3 are between those of Case 1 and Case 4. The network state of the two cases will gradually converge to a final state which is between 0 and 1.

In Figure 2, the four continuous curves and four groups of discrete points represent the population dynamics derived by the theory and simulation, respectively. It can be seen that the simulations agree with the theoretical results. The small gaps are due to the ignorance of the dependence between network states and network degree. In all four cases, the global network states all converge to the stable state, known as the ESSs.
When the population dynamics equals to zero, the whole network reaches a stable state. Therefore, making equations (1) and (2) equal to 0, one can get the expression of ESSs of both uniform and non-uniform degree networks, which are shown below (Jiang et al., 2014a, 2014b).

**Theorem 3:** In an $N$-users 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 rules, the evolutionary stable network states can be summarized as follows:

\[
\hat{p}_f^* = \begin{cases} 
1, & \text{if } u_{ff} > u_{fn} > u_{nn} \\
0, & \text{if } u_{nn} > u_{fn} > u_{ff} \\
\frac{(k-2)u_{fn} + u_{ff} - (k-1)u_{nn}}{2u_{fn} - (k-2)(u_{ff} - u_{nn})}, & \text{else}
\end{cases}
\]  

Similarly, for non-uniform degree network, we can obtain the following theorem (Jiang et al., 2014a, 2014b).

**Theorem 4:** In an $N$-users 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:

\[
\hat{p}_f^* = \begin{cases} 
1, & \text{if } u_{ff} > u_{fn} > u_{nn}, \\
0, & \text{if } u_{nn} > u_{fn} > u_{ff}, \\
\frac{\left(\frac{k^2}{k-2}\right)(u_{fn} - u_{nn}) + (u_{ff} - u_{nn})}{\left(\frac{k^2}{k-2}\right)\left(2u_{fn} - u_{ff} - u_{nn}\right)}, & \text{else}
\end{cases}
\]

The results are consistent with the intuition. If one can get more utility from forwarding information regardless of the neighbors’ strategies, he will definitely forward this information, i.e. the probability of forwarding is one in the steady state. Similarly, a steady
loss will make nobody want to forward information, which is described by $p^* = 0$. For the situations between two extremities above, the ESS will be determined by the exact value of $u_{ff}$, $u_{fn}$ and $u_{nn}$, and one can have a general impression through the simulation results as shown below.

Figure 3 illustrates the graph structure and the degree characteristics of a Facebook subnetwork, which consists of ten subgraphs. With the payoff matrix given above, the simulation results and theoretical results of the ten subgraphs are shown in Figure 4. The horizontal red line and the cyan line show the ESSs of Case 1 and Case 4, while the green one and the brown one show the ESSs of Case 2 and Case 3. The gaps between simulation and theoretical results of Case 2 and Case 3 are from the neglected dependence between the network state and the network degree. From the results, it can be seen that the theory is consistent with the simulation.
The theorems above characterize the social networks with the same type of users. The study of the information diffusion over heterogeneous social network is also conducted (Cao et al., 2016). Two different situations need to be considered: the unknown user type model and the known user type model. For the unknown user type model, each user does not know the exact type of the neighbors and tends to consider them the same as itself. For the known user type model, each user knows the exact type of the neighbors and thus their decisions.

For the unknown user type model over heterogeneous social network, the following theorem can be derived (Cao et al., 2016).

**Theorem 5:** (Evolutionary dynamics): In the unknown user type model, the evolutionary dynamics for the network states $p_f(i)$ and $p_f$ are given in the equations as follows:

$$
\dot{p}_f(i) = \frac{1}{N} p_f - \frac{1}{N} p_f(i) + \frac{\alpha}{N} p_f (p_f - 1) \\
\times \Delta(i) \left( \left( \bar{k} - 3 + \frac{2}{k-1} \right) p_f + 1 - \frac{1}{k-1} \right) + \Delta_n(i) \left( \bar{k} - 1 \right),
$$

$$
\dot{p}_f = \frac{\alpha}{N} p_f (p_f - 1) \left[ \bar{\Delta} \left( \left( \bar{k} - 3 + \frac{2}{k-1} \right) p_f + 1 - \frac{1}{k-1} \right) + \bar{\Delta}_n \left( \bar{k} - 1 \right) \right],
$$

where $\bar{\Delta} := \sum_{i=1}^{M} q(i) \Delta(i)$ and $\bar{\Delta}_n := \sum_{i=1}^{M} q(i) \Delta_n(i)$.

**Theorem 6:** (ESSs): In the unknown user type model, the ESSs of the network are as follows:

$$
p_f^* = \begin{cases} 
0, & \text{if } \bar{u}_{nn} > \bar{u}_{fn} \\
1, & \text{if } \bar{u}_{ff} < \bar{u}_{fn} \\
\frac{\bar{\Delta}_n(1 - \bar{k}) + \bar{\Delta}(\bar{k} - 1 - 1)}{\Delta(\bar{k} - 3 + 2\bar{k}^{-1})}, & \text{if } \max\{\bar{u}_{ff}, \bar{u}_{fn}\} < \bar{u}_{fn},
\end{cases}
$$

$$
p_f^*(i) = p_f^* + \alpha p_f^* (p_f^* - 1) \left[ \Delta(i) \left( \left( \bar{k} - 3 + \frac{2}{k-1} \right) p_f^* + 1 - \frac{1}{k-1} \right) + \Delta_n(i) \left( \bar{k} - 1 \right) \right],
$$

where $\bar{u}_{ff} = \sum_{i=1}^{M} q(i) u_{ff}(i)$ and $\bar{u}_{fn}$, $\bar{u}_{nn}$ are similarly defined.

For the known user type model, the following theorems can be obtained (Cao et al., 2016).

**Theorem 7:** In the known user type model, the population dynamics $p_f(i)$ and the relationship dynamics $p_{ff}(i, l)$ for $i \neq l$ and $i = l$ are given as follows:
\[
\dot{p}_f(i) \approx \frac{\alpha - k}{N} p_f(i) p_u(i, j) (p_{ui}(i, j) + p_{uf}(i, j)) \\
\times \sum_{j=1}^{M} q(j) [p_{uf}(j, i) u_{uf}(i, j) + p_{uf}(j, i) u_{un}(i, j)] - p_{uf}(j, i) u_{uf}(i, j) - p_{uf}(j, i) u_{un}(i, j),
\]

\( (9) \)

\[
\dot{p}_{ff}(i, l) = \frac{2}{N} q(i) q(j) p_f(i) (1 - p_{ff}(i, j)) \times \left[ \frac{p_f(l)}{p_n(i)} (1 - p_{ff}(i, l)) - p_{ff}(i, l) \right] \\
+ \frac{2}{N} q(i) q(l) p_f(l) (1 - p_{ff}(l, i)) \times \left[ \frac{p_f(i)}{p_n(l)} (1 - p_{ff}(l, i)) - p_{ff}(i, l) \right],
\]

\( (10) \)

\[
\dot{p}_{ff}(i, l) = \frac{2}{N p_n(i)} q^2(i) p_f(i) (1 - p_{ff}(i, j)) (p_f(i) - p_{ff}(i, j)).
\]

\( (11) \)

Figure 5 shows the simulation results of the population dynamics for unknown user type model, where it can be seen that theoretical results are consistent with the simulation results. The small gap is probably because of the ignorance of higher orders of \( \alpha \). These results are based on the parameters: \( u_{ff}(1) = 0.4, u_{ff}(2) = 0.2, u_{un}(1) = 0.6, u_{un}(2) = 0.4, u_{nn}(1) = 0.3, u_{nn}(2) = 0.5, N = 1000, k = 20, q(1) = q(2) = 0.5, \alpha = 0.05. \)
Figure 6 shows the results for the known user type model. As one can see, the population dynamics of known type users make a very large difference than those of unknown type users, which indicates the necessity of type classification. In this simulation, the payoff matrix is chosen as follows with other parameters $N = 1,000$, $k = 20$, $q(1) = 0.5518$, $q(2) = 0.4482$, $\alpha = 0.05$.

$$
u_{ff} = \begin{bmatrix} 0.5882 & 0.0116 \\ 0.8688 & 0.1590 \end{bmatrix}, \quad 
u_{fn} = \begin{bmatrix} 0.9619 & 0.7370 \\ 0.5595 & 0.7180 \end{bmatrix},$$

$$
u_{nf} = \begin{bmatrix} 0.9339 & 0.9864 \\ 0.3288 & 0.4593 \end{bmatrix}, \quad 
u_{nn} = \begin{bmatrix} 0.2479 & 0.3385 \\ 0.6570 & 0.2437 \end{bmatrix},$$

Notes: The heterogeneous game theoretic model can predict future diffusion dynamics (a) using data up to Time 22; (b) using data up to Time 41.
The prediction of the future diffusion dynamics is also illustrated in Figure 7. It can be seen that the graphical EGT can well predict the diffusion dynamics in the future based on past data.

6. Conclusion
In this paper, we summarized three papers (Jiang et al., 2014a, 2014b and Cao et al., 2016) to discuss how to use the graphical evolutionary game theory to analyze the static state and the dynamics of information diffusion in social network.

References
Kimura, M., Saito, K. and Motoda, H. (2009), “Blocking links to minimize contamination spread in a social network”, ACM Transactions on Knowledge Discovery from Data (Data), Vol. 3 No. 2, pp. 91-9:23.


Corresponding author
Benliu Qiu can be contacted at: 2016020902016@std.uestc.edu.cn

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com