Modeling and Simulating the Empathetic Decision-Making Behavior in Social Networks with Homophily

Chien-Min Chen, Min-Sheng Lin

Department of Electrical Engineering, National Taipei University of Technology, Taiwan
{t5319011, mslin}@ntut.edu.tw

Abstract

People tend to associate with other people whose qualities are similar to their own. This phenomenon is the so-called "like attracting like", or the "homophily", in social networks. This paper presents a model that simulates how individuals in social networks with homophily make decisions as a result of the empathic effect, which refers to influence by the decisions of friends or the populations of social networks of which they are a member. In this study, experiments were conducted using Monte Carlo simulation to test a particular decision-making preference (such as asking a question in a Facebook group) and Chi-square tests were conducted to see whether the distribution of people's final choices is the same as that of people's intrinsic choices. Experimental results show that under each the following conditions, individuals are less likely to experience the empathic effect, meaning that they are less likely to change their intrinsic choices: (1) the social network exhibits a more homophily; (2) decision-makers have more options from which to choose, and (3) individuals have fewer friends.

Keywords: Simulation model, Social networks, Empathetic decision-making, Homophily, Small-world network

1 Introduction

A social network is a set of people who are connected by social relationships of one or more types and who share common interests. People use social networks to see what their friends are doing; show them that they care about them, and obtain new information from the messages that their friends share directly on such networks.

One of the most basic concepts concerning the structure of social networks is *homophily*, which refers to the fact that people tend to be similar to their friends [1]. When people have similar personalities and preferences, they are likely to please each other and exhibit the same decision-making behaviors. This phenomenon underlies "sharing vile habits" or the fact that "birds of a feather flock together", which relate to a similarity between attributes [2]. When social interactions that are based on homophily are established, friends are more likely to maintain contact with each other.

In an environment with greater homophily, decision-makers are more likely to put themselves in other people's shoes; consider their psychological reactions, and understand their attitudes and emotional competences before making judgments and decisions concerning events. This phenomenon is called *empathy* or "psychic mobility". Empathy easily occurs among individuals with various backgrounds. When emotions rise to a certain level that inhibits objective judgment, decision-makers begin to empathize with other decision-makers as if they were important figures in their lives. Empathy typically occurs when decision-makers inadvertently do or say something that triggers unsolved problems that occupy others' minds. These problems typically concern families, including parents and siblings, as well as friends and other important figures.

This study examines which factors, such as population size, mean number of friends per individual, homophily level, social network structure, and the number of available options in a decision, affect the empathetic decision-making behaviors of individuals in social networks and how they do so.

The main contributions of this paper are as follows. First, a model of social networks with various homophily levels is developed. Second, a computerized mathematical model of empathetic decision-making behavior that considers intrinsic preferences and empathetic preferences is proposed. Third, experiments are conducted using Monte Carlo simulation to test the hypothesis that the distribution of people's final choices is the same as that of their intrinsic choices.

The rest of the paper is organized as follows. Section 2 provides a brief overview of relevant literature. Section 3 describe the methodology that is used to model empathetic decision-making behavior in social networks with homophily. Section 4 explains the experimental setup and Section 5 discusses the experimental results. Section 6 draws conclusions.

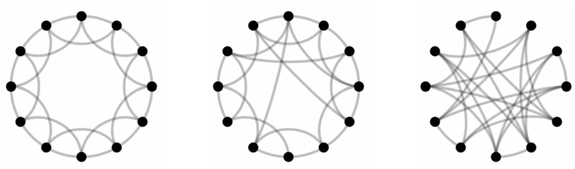
*Corresponding Author: Chien-Min Chen; E-mail: t5319011@ntut.edu.tw
DOI: 10.3966/160792642018031902009

2 Literature Review

The reviews of literatures are organized as follows. First, the small-world network that is used to model the proposed social network in Subsection 3.1 is studied. Second, the literature on the concept of homophily, which will be incorporated into the small-world network in Subsection 3.2, is reviewed. Finally, the literature most closely related to this paper and, in particular, research on the empathetic decision-making behavior in social networks, which will be studied in Subsection 3.3, is briefly discussed.

2.1 Small-World Network

In 1976, Milgram [3] was the first to raise the small-world issue in interpersonal relationships. He performed a series of experiments to confirm that the distance between people is not as far as imagined and only six degrees of separation (people) separate everyone in the world. The small-world network model can be generated using a defined regular graph and a probability of rewiring the links p_r. When the rewiring probability $p_r = 0$, no irregular connections occur and a regular graph is formed, in which every node is connected to several neighbor nodes and the clustering is maximal; when the random link probability $p_r = 1$, all of the regular connections become irregular and a random graph is formed; this graph has the smallest average path length. The results of the experiment show that when the rewiring probability p_r is between 0 and 0.15, the average path length decreases rapidly while clustering decreases slowly; this finding which consistent with the fact that the small-world network model has a small average path length and a large average clustering coefficient [4]. This situation is schematically in Figure 1 [5].



$p_r = 0$ $p_r = 0.15$ $p_r = 1$

Figure 1. Diagram Showing Changes in Regular Network – Small-World Network – Random Network [5]

A variety of social networks, particularly friendship networks, have been shown to exhibit the small-world phenomenon. One such example was considered by Wohlgemuth and Matache [6], who showed that Facebook group networks have small average path lengths and large clustering coefficients that do not vanish as the network size increases; they therefore exhibit small-world features. This study will use the small-world network as the default social network topology.

2.2 Homophily

McPherson *et al.* [7] showed that individuals with similar characteristics, such as gender, race, ethnicity, age, and educational history, tend to be friends; this phenomenon is called homophily. The homophily phenomenon has been widely studied extensively in various domains. Aral *et al.* [8] developed a dynamic matched sample estimation framework to distinguish influence and homophily effects in dynamic networks. Their results showed that previously developed methods significantly overestimated peer influence in such networks, mistakenly identifying homophily-driven diffusion as influence-driven contagion. Yavaş and Yücel [9] identified the conditions under homophily reinforces or undermines diffusions over social networks. Small increases in homophily favor diffusion whereas large increases disfavor it. Gerwen and Sleeuwen [10] examined the differences between opportunity-based and choice-based homogeneities within subgroups, based on the gender and socioeconomic status of 1640 university students from one university in the U.S. Their results revealed that Facebook networks of individuals with a higher socioeconomic status are more homogeneous with respect to major field of study than are Facebook networks of individuals with a lower socioeconomic status.

Pin and Rogers [11] described the process of formation of a social network with homophily as having two stages. First, the characteristic of a node is chosen based on some type of homophily; second, nodes form connections with a probability that increases with the similarity of this characteristic. To analyze the empathetic effect at various homophily levels, Subsection 3.2 proposes a novel approach to generate a social network with homophily.

2.3 Empathetic Decision-making Behavior

One research study that is closely related to this work was performed by Salehi-Abari and Boutilier [12]. The authors proposed a model of empathetic social choice in which individuals derive utility based on both their own intrinsic preferences and their empathetic preferences that are derived from their friends. The authors then converted the problem into a weighted form of classical preference aggregation of which social welfare maximization and certain forms of voting are examples. Many significant differences exist between our work and that of Salehi-Abari and Boutilier [12]. First, their empathetic model assumed that person's intrinsic preferences are not correlated with those of their friends. However, this work proposes a novel model for generating a social network that exhibits homophily, in which individuals were likely to have the same intrinsic preferences as their friends. Second, they performed their experiments based on the impartial culture assumption which requires that, given a number of alternatives, each individual assigns a random and equal probability (uniform distribution model) to all possible preference rankings. Since the uniform distribution model is not very realistic, this work provides experimental results concerning some other distribution models, such as truncated-normal and Zipf distribution models. Third, a statistical (hypothesis testing) method is used herein to analyze the effects of various factors on an individual's empathetic decision-making behavior; however, Salehi-Abari and Boutilier [12] used the index of social welfare loss as a performance metric in various empathetic models.

3 Modeling Empathetic Decision-Making Behavior in Social Networks with Homophily

This section introduces the use of small-world networks to create a model of social networks, and then defines homophily level in relation to people's intrinsic choices in a particular decision-making preference. It finally presents a model of empathetic decision-making behavior and discusses how individuals make final choices.

3.1 Modeling a Social Network as a Small-world Network

The simulation model in this research is based on a small-world network. Methods for establishing the small-world network from the literature [5] are referred to. These methods firstly generate a regular graph and then, based on the rewiring probability p_r, select some links from the graph to be re-linked to non-neighboring nodes. The steps of constructing a small-world network with n nodes and a mean degree k are described in detail below.

Algorithm 1: ***Modeling social network based on small-world network***

Step 1: Place n nodes in a ring lattice and label the nodes with numbers from 1 to n according to their respective positions around the ring.

Step 2. Connect every node to k neighbors, $k/2$ on either side.

Step 3: Randomly rewire every link with rewiring probability p_r so that self-connections and duplicate links are not formed.

end algorithm

Table 1 presents some basic properties of small-world networks (n = 4000, k = 50) that are rewired with probability p_r = 0.05, 0.15 and 1 in the experiments, and Figure 2 displays the distributions of the degrees of the nodes in these small-world networks.

Table 1. Basic properties of small-world networks (n = 4000, k = 50)

rewiring prob. (p_r)	avg. path length $L(p_r)$	avg. clustering coefficient $C(p_r)$	network diameter $D(p_r)$
0.05	3.077	0.631	5
0.15	2.809	0.454	4
1.00	2.517	0.012	3

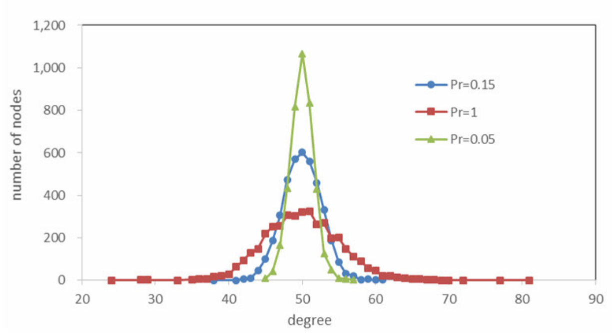


Figure 2. Distributions of degrees of nodes in small-world networks with various rewiring probabilities p_r (n = 4000, k = 50)

3.2 Modeling Homophily in Social Networks

People are typically more likely to build social links with people who are similar to them. This phenomenon is called homophily: "birds of a feather flock together". Homophily is evident in connections in social networks, such as marriage, friendship, connection with co-workers, and co-membership, for example, as can be evidenced in interests, education, political views and other characteristics. People commonly consider differences between traits and those of others before deciding whether they want to build social links with others. Therefore, people with the same traits are typically more likely to become friends, and individuals and members of their social networks generally have homogeneous characteristics. Since the homophily principle limits the diversity in how people obtain information, and affects how individuals perceive certain things and make decisions, the homophily level among individuals in social networks must be considered in elucidating how individuals make decisions in social networks.

Here, homophily among intrinsic choices made by individuals at a particular preference is considered. For example, the following survey question appears on Facebook; "Which do you prefer, baseball or football?" People in Taiwan and most of their friends will probably answer "baseball" while those in Brazil are more likely to answer "football". Therefore, people's intrinsic choices concerning this question exhibit high homophily. In contrast, if people's intrinsic choices concerning a particular preference have low homophily, then the inconsistency between their and their friends' intrinsic choices will be greater.

For example, the following survey question appears on Facebook; "What is your favorite color?" Since the color preferences of individuals typically do not depend on their educational backgrounds, religion or places of residence, people in Taiwan and their friends will probably not all prefer a certain color; the same will be true for people in Brazil. Restated, people's intrinsic choices with respect to the questoin, "What is your favorite color?" have a low homophily level. Based on this concept, it is necessary to specify individuals' intrinsic choices related to the homophily level with repect to a particular preference before information about their final choices begins to diffuse across a social network.

The complete sequence of steps in the proposed approach for modeling homophily in a social network is as follows.

Step 1. Let $P = \{1, 2, ..., n\}$ denote the set of nodes (individuals) in a social network and let $O = \{1, 2, ..., m\}$ denote the set of available options to be chosen with respect to a particular preference. First, a small-world network with n nodes (individuals) and mean degree k is generated using the algorithm in Subsection 3.1. In the following steps, let $a_{i,x} = 1$ if the intrinsic choice that is made by individual i is x; otherwise $a_{i,x} = 0$. Initially, assume $a_{i,x} = 0$ for all individuals $i \in P$ and all options $x \in O$.

Step 2. Assume that X is as a random variable (which may follow a uniform distribution, truncated-normal distribution, or Zipf distribution, for example) that is used to generate intrinsic choices. Use X to generate n random numbers $(x_1, x_2, \ldots, x_n)$, $x_i \in O$, for $1 \le i \le n$. Then, $(x_1, x_2, \ldots, x_n)$ are reordered to $(x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)})$ such that $x_{\pi(1)} \le x_{\pi(2)} \le \ldots \le x_{\pi(n)}$. Accordingly, identical options $x_{\pi(i)}$ are placed in consecutive positions. Then, for all individuals $i \in P$, let $a_{i,x_{\pi(i)}} = 1$, meaning that the intrinsic choice of individual i is $x_{\pi(i)}$. Since the small-world network that was constructed in Step 1 is characterized by high clustering, the probability that node i connects with neighboring nodes $i-1$, $i+1$, $i-2$, $i+2$, …, $i-k/2$, $i+k/2$ (meaning that they become friends with each other) is high. Therefore, the probability that an individual and his or her friends make the same intrinsic choice is high because identical options are placed in consecutive positions, as mentioned above. Restated, this social network has a maximum homophily level.

Step 3. Next, adjustments are made to the homophily in this social network based on the selected homophily level $p_h \in [0..1]$. Each individual is assumed to have probability $1-p_h$ of generating a new random value x^* from the random variable X and his or her intrinsic choice will then be reset to x^*. If p_h equals 1, then the intrinsic choices that are made by all the members in this social networks are not reset and these intrinsic choices have the highest homophily level, which was obtained in Step 2; if p_h equals 0, then all the individuals will reset their intrinsic choices based on the random variable X and these intrinsic choices have the lowest homophily level.

Based on the above description, the formal algorithm for modeling the homophily in a social network is presented below.

Algorithm 2: ***Modeling homophily in social networks***

Step 1: Create a social network according to Subsection 3.1; for **each individual** $\boldsymbol{i \in P}$ and each option $x \in O$ **do** $a_{i,x} \leftarrow 0$; // initially, reset all $a_{i,x}$ to zero//
end for

Step 2: Use a specific random variable X to generate n random numbers $(x_1, x_2, \ldots, x_n)$ with values that range from 1 to m; Reorder $(x_1, x_2, \ldots, x_n)$ to $(x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)})$ such that $x_{\pi(1)} \le x_{\pi(2)} \le \ldots \le x_{\pi(n)}$;
for each individual $i \in P$ **do** $a_{i,x_{\pi(i)}} \leftarrow \mathbf{1}$**;**
end for

Step 3: **for** each individual $i \in P$ **do**
Use uniform distribution $U[0,1]$ to generate a random number r;
if $r \ge p_h$ **then**
Use the specific random variable X to generate a random number x^*;
$a_{i,x_{\pi(i)}} \leftarrow 0$; $a_{i,x^*} \leftarrow 1$;
end if
end for

end algorithm

3.3 Modeling Empathetic Decision-making Behavior in Social Networks

How do individuals in social networks make decisions when they encounter a choice and have many available options to choose from? They consider intrinsic preferences (which are related to intrinsic choices, defined in the previous subsection), but they are also influenced by the decisions that are made by their friends or fellow members of their netwrok. Such influence is called empathetic preference. For example, when an individual fills out a questionnaire in a Facebook group, he or she can see statistics concerning available selected options, including statistics concerning the choices that were made by his or her friends. These statistics may affect the decision of the individual, causing him or her to change his or her intrinsic choice. Accordingly, the model of empathetic decision-making in social networks is as follows.

Step 1. Let P and O denote the set of nodes (individuals) in the social network and the set of available options that can be chosen in relation to a particular preference, respectively. Use the algorithm in Subsection 3.2 to specify the intrinsic choices of all individuals i such that $a_{i,x} = 1$ if the intrinsic choice that is made by individual i is x; otherwise $a_{i,x} = 0$. In the following steps, let $b_{i,x} = 1$ if the final choice that is made by individual i is x; otherwise $b_{i,x} = 0$. Initially, assume $b_{i,x} = 0$ for all individuals $i \in P$ and all options $x \in O$.

Step 2. Randomly select s individuals from P as initial seeds of S and their final choices are assumed to be entirely based on their own intrinsic choices rather than being affected by their friends or others, so $b_{i,x}$ is set to $a_{i,x}$ for $i \in S$, $x \in O$.

Step 3. Let P^* denote the subset of individuals in the population that has made final decisions and let $F(i)$ represent the set of friends of individual i. Assume that w_o represents the strength of an individual's own intrinsic preference; w_f is the strength of the individual's empathetic preference that is derived from influence by friends, and w_p is the strength of empathetic preference that is derived from the influence of population of the social network. Define $w(i,x)$ as the strength of the preference of individual i for option x, which is a linear combination of the abovementioned three preference as follows.

$$w(i,x) = w_o \cdot a_{i,x} + \sum_{j \in F(i) \cap P^*} w_f \cdot b_{j,x} + \sum_{j \in P^*} w_p \cdot b_{j,x}\ ,$$

for $i \in P-P^*$, $x \in O$.

According to the above equation, individual i may choose option x_i^* with the maximum value $w(i,x)$ as his or her final choice and then b_{i,x_i^*} is set to one. The value of x_i^* is calculated as follows.

$$x_i^* = \arg \max_{x \in O} \{w(i,x)\}\ , \text{ for } i \in P-P^*.$$

Based on the above two equations for the calculation of $w(i,x)$ and x_i^*, the final choice of individual i will related to the choices of his or her friends and the population P^* of the social network that has made final decisions. Notably, the time when individual i receives information from his or her friends will affect his or her final decision. The probability that individual i receives this information is assumed to depend on the number of friends who have made their decisions ($|F(i) \cap P^*|$). Basically, a greater $|F(i) \cap P^*|$ value indicates that individual i is more likely to receive the information soon and become the next decision-maker.

The formal algorithm for modeling the empathetic decision-making process in social networks is as follows.

Algorithm 3: ***Modeling empathetic decision-making in social networks***

Step 1: According to Subsection 3.2, specify the intrinsic choice $a_{i,x}$ for all individuals $i \in P$ and all options $x \in O$; **for** each individual $i \in P$ **and** each option $x \in O$ **do** $b_{i,x} \leftarrow 0$; // initially, reset all $b_{i,x}$ to zero //
end for

Step 2: Randomly select s individuals as initial seeds of S; **for** each individual $i \in S$ **and** each option $x \in O$ **do** $b_{i,x} \leftarrow a_{i,x}$;
end for

Step 3: $P^* \leftarrow S$;
while $P^* \subset P$ **do**
Randomly select individual i from $P-P^*$ with the probability proportional to $|F(i) \cap P^*|$;
for each option $x \in O$ **do**
$w(i,x) = w_o \cdot a_{i,x} + \sum_{j \in F(i) \cap P^*} w_f \cdot b_{j,x} + \sum_{j \in P^*} w_p \cdot b_{j,x}$,
end for
$x_i^* \leftarrow \arg \max_{x \in O} \{w(i,x)\}$;
$b_{i,x_i^*} \leftarrow 1$;
$P^* \leftarrow P^* \cup \{i\}$;
end while

end algorithm

4 Simulation Design and Methods

A simulation experiment is conducted using the model that was presented in Section 3 and the Monte Carlo method. Random samples are generated using random numbers. The same experiment is repeated many times. Finally, the experimental results are analyzed using statistical methods. The simulation platform is built using the NetLogo, which is a very popular multi-agent simulation tool and that is frequently used to simulate various behavioral patterns in social networks.

4.1 Simulation Parameters

The user-friendly interface that is provided by NetLogo enables easy adjustment of various experimental parameters. Figure 3 shows a snapshot of the NetLogo simulation platform, which allows different combinations of parameters to be input to generate different outputs. As shown in Figure 3, the platform enables the adjustment of many parameters. Table 2 defines the more important parameters in the proposed model.

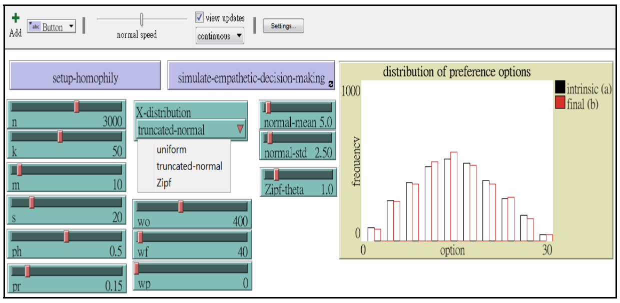


Figure 3. Snapshot of NetLogo interface of simulation platform

Table 2. Summary of parameters used in the simulation model

Parameter	Description	Possible values
n	population size (number of nodes)	1000-10000
k	mean number of friends per individual (mean degree of nodes)	20-200
m	number of available options for decision making	2-50
s	number of initial seeds	100
p_h	homophily level	0-1
p_r	rewiring probability in the small-world network	0-1
w_o	strength of an individual's own intrinsic preference	1800
w_f	strength of empathetic preference derived from the influence of friends	200-1000
w_p	strength of empathetic preference derived from the influence of the social network population	0-10
X	distribution of intrinsic choices for each of the available options	Uniform, T-normal, Zipf
θ	Zipf rank exponent	0-1

Parameter X represents a random variable to generate the intrinsic choices that were described in Subsection 3.2. In this experiment, simulations are performed using the three probability models - uniform, truncated-normal (T-normal) and Zipf. The following figure shows the three probability distribution functions that were generated by NetLogo with n = 3000 and m = 10.

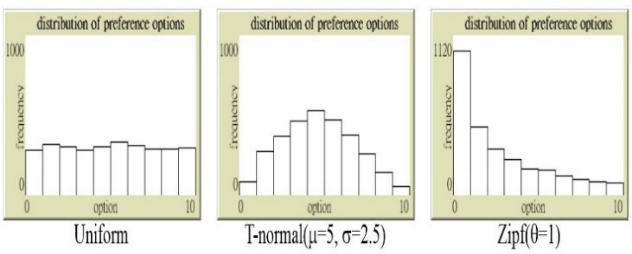


Figure 4. Probability distributions of intrinsic choices used in simulation model

Given that $f(x)$ represents the probability that the xth option is randomly among m options, the following subsections describe the three probability models in detail.

Uniform model. The probability distribution function is:

$$f(x) = \frac{1}{m}.$$

The uniform distribution specifies that the each option has the same probability of being chosen. For example, a tossed coin is equally likely to land head or tail side up, just as a thrown die is equally likely to land with one of its six sides' facing upward.

Truncated-normal (T-normal) model. The normal distribution is commonly used to describe some statistics that arise in biology and nature. For example, the distribution of body heights in an ethnic group and that of air temperatures in an area are both normally distributed. The T-normal distribution is similar to the normal distribution with the difference that the values must be between the given values c_1 and c_2. If x has a normal distribution with mean μ and standard deviation σ, then the density of the T-normal distribution is

$$f(x|c_1 \le x \le c_2) = \frac{\frac{1}{\sigma}\varphi(\frac{x-\mu}{\sigma})}{\Phi(\frac{c_2-\mu}{\sigma}) - \Phi(\frac{c_1-\mu}{\sigma})},$$

where φ and Φ are the density and distribution functions of the standard normal distribution.

The above equation indicates that the probability distribution for $f(x|\ c_1 \le x \le c_2)$ is determined by the four parameters μ, σ, c_1 and c_2. Since the x boundary in this simulation experiment is apparently [1, m], c_1 and c_2 are set to 1 and m, respectively; to simplify the experiment, the mean μ and standard deviation σ are set to m/2 and m/4, respectively, so that they form a symmetric T-normal distribution. Since the T-normal distribution is a continuous random variable, rounding is performed to obtain discrete integer values.

Zipf model. In the Zipf model, options are numbered in descending order based on $f(x)$ probability: $f(x = 1) \ge f(x = 2) \ge \ldots \ge f(x = m)$. The probability distribution function is,

$$f(x) = \frac{1}{x^{\theta}\sum_{j=1}^{m}(\frac{1}{j})^{\theta}},$$

where θ is called the Zipf rank exponent. Zipf's law states that the jth most frequently appearing object will appear $1/j^{\theta}$ times as often as the most frequently appearing object. The Zipf distribution has a wide range of applications. For example, the frequency of an English word is inversely proportional to its rank in the frequency table based on the logarithm of its frequency and the same is true for the population of a city to its rank in the population table.

4.2 Evaluation Metrics

This simulation experiment mainly examines how various parameters affect the empathetic decision-making behavior of individuals in social networks. For example, are the results (final choice statistics) that are obtained by asking a question in a Facebook group the same as those (intrinsic choice statistics) obtained by conducting a traditional telephone survey? Statistical testing techniques are used to test the results of the experiment. The null hypothesis H_0 to be tested in this experiment is as follows.

H_0: *The distribution of people's final choices is the same as that of their intrinsic choices*.

This hypothesis is tested using the Chi-squared goodness-of-fit test. Assume that A_j and B_j represent the numbers of people who choose the jth option as their intrinsic choices and those who choose the same as their final choices, respectively. The equation for the test statistic χ^2 is as follows.

$$\chi^2 = \sum_{j=1}^{m} \frac{(B_j - A_j)^2}{A_j}$$

Common approaches to testing H_0 are the critical value approach and the p-value approach. The latter is used in this work. A p-value is the probability that an observed (or more extreme) result arises by chance if the null hypothesis is true. Since the above χ^2 has a Chi-square distribution with m–1 degrees of freedom, the corresponding p-value can be calcultated using statistical software. If the p-value is less than the significance level α, then H_0 is rejected; in contrast, if the p-value is greater than the significance level α, then H_0 is not rejected. The value of the significance level α is typically set to 0.01, 0.05 or 0.1.

The p-value is used as an evaluation metric for our simulation results and the p-values that are generated using various combinations of parameters can be compared. A smaller p-value means that an individual is more likely to be affected by his or her friends or fellow members, such that the empathetic effect changes his or her decision; in contrast, a greater p-value means that an individual is less likely to experience the empathetic effect and his or her final choice is more likely to be the same as his or her intrinsic choice.

5 Experimental Results

This study provides insights by examining the following questions through a series of experiments.

- What is the effect of population size on p-value?
- What is the effect of mean number of friends per individual on p-value?
- What is the effect of homophily level on p-value?
- What is the effect of number of available options in decision-making on p-value?
- What is the effect of social network structure on p-value?
- What is the effect of strength of the empathetic preference on p-value?
- What is the effect of Zipf rank exponent on p-value?

The experiments herein are performed according to the design in Section 4 using three distributions of intrinsic choices to generate each of the available options - Uniform, T-normal and Zipf. Each experiment is repteated many times with the same values of the parameters, and the average p-values are reported to answer the above questions.

5.1 Effect of Population Size

Figure 5 presents the experimental results. Regardless of the population of the social network, the p-values in the three distributions of intrinsic choices follow the order Uniform model > T-normal model > Zipf (θ = 0.6) model. When the significance level α is set to 0.1, the p-values in both the Uniform and T-normal models are greater than α (= 0.1), meaning "fail to reject the null hypothesis H_0", so the distribution of people's final choices is the same as that of their intrinsic choices in these two models. However, the p-value in the Zipf model decreases as the population increases. When p-value < α (= 0.1) (the population is greater than 3000), this suggests "reject the null hypothesis H_0", and the distribution of people's final choices is different from that of their intrinsic choices when the population is greater than 3000 in the Zipf model.

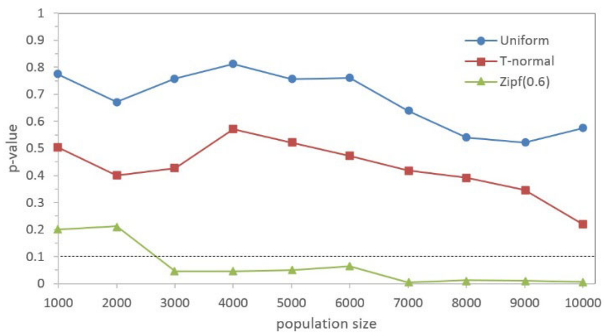


Figure 5. Correlation between P-value and population size (k = 50, m = 10, s = 100, p_h = 0.5, p_r = 0.15, w_o = 1800, w_f = 200, w_p = 0, θ = 0.6)

5.2 Effect of Mean Number of Friends per Individual

Figure 6 shows that a larger mean degree per node (a larger mean number of friends per individual) leads to a lower p-value, so an individual with more friends is more likely to be influenced by them and thus to change his or her intrinsic choice. For example, when the mean number of friends per individual exceeds 160 in the Uniform model or 120 in the T-normal model, the p-value < α (= 0.1), suggesting "reject the null hypothesis H_0" and the distribution of people's final choices differs from that of their intrinsic choices.

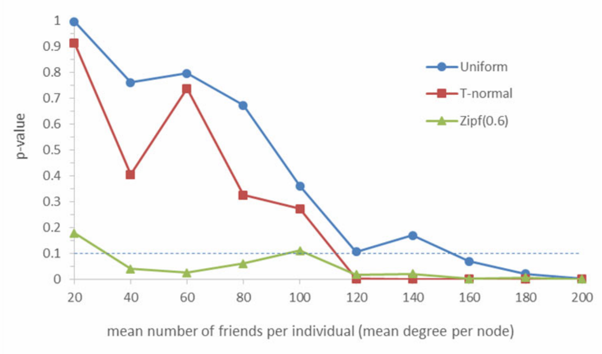


Figure 6. Correlation between P-value and mean number of friends per individual (n = 4000, m = 10, s = 100, p_h = 0.5, p_r = 0.15, w_o = 1800, w_f = 200, w_p = 0, θ = 0.6)

5.3 Effect of Homophily Level

Figure 7 reveals that greater homophily results in a greater p-value, so when the homophily level among the intrinsic choices of an individual and his or her friends is higher, the distribution of people's final choices is more likely the same as that of their intrinsic choices. The experiments herein show that when the homophily level > 0.5, "fail to reject the null hypothesis H_0" for the Uniform model, the T-normal model and the Zipf model.

Notably, when the homophily level < 0.2, the Uniform model has a large p-value because when the homophily level is insignificant, the intrinsic choices that are made by an individual's friends are almost randomly distributed in the Uniform model and the number of those who choose any particular option is not significant. Therefore, the individual is less likely to experience the empathetic effect and change his or her decision. However, as the homophily level increases, the number of an individual's friends who choose a particular option also increases, causing the individual to become more susceptible to the empathetic effect and to more likely to change his or her decision accordingly. Nevertheless, when the homophily reaches a certain level (0.2 in this experiment), the probability that the distribution of people's final choices is the same as that of people's intrinsic choices begins to increase; the individual becomes actually less likely to change his or her decision under the influence of friends and the p-value gradually increases.

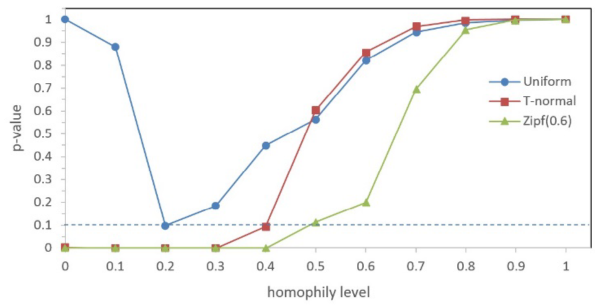


Figure 7. Correlation between P-value and homophily level (n = 4000, k = 50, m = 10, s = 100, p_r = 0.15, w_o = 1800, w_f = 200, w_p = 0, θ = 0.6)

5.4 Effect of Number of Available Options in Decision-Making

Figure 8 demonstrates that more available options yield a larger p-value, so an individual is less likely to experience the empathetic effect when he or she has more options in making a decision, and vice versa. For example, when the number of options is less than five in the T-normal model or less than 15 in the Zipf model, the p-value < α (= 0.1), suggesting "reject the null hypothesis H_0" so the distribution of people's final choices differs from that of their intrinsic choices. Accordingly, surveys via social networks may produce different statistical results from traditional surveys when few options are available.

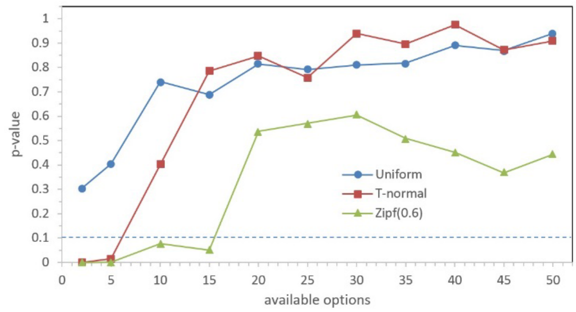


Figure 8. Correlation between P-value and number of available options (n = 4000, k = 50, s = 100, p_h = 0.5, p_r = 0.15, w_o = 1800, w_f = 200, w_p = 0, θ = 0.6)

5.5 Effect of Social Network Structure

This experiment examines whether the structure of a social network affects individual decision-making. The first step is to eliminate any effect of homophily by setting its level to zero. Next, the rewiring probability of the small-world network, p_r, is adjusted to alter the small-world network structure. Figure 9 presents the experimental results. The right-hand axis represents the ratios $L(p_r)/L(0)$ and $C(p_r)/C(0)$ for various p_r, where L(0) and C(0) are the average path length and average clustering coefficient in the regular network (when p_r = 0), respectively, and $L(p_r)$ and $C(p_r)$ are the same quantities in the small-world network rewired with probability p_r.

In the Uniform model, the T-normal model, and the Zipf model, the p-value is almost independent of $L(p_r)$ and $C(p_r)$, which are the two characteristics of the social network structure. Therefore, under the assumption of zero homophily in the social network and given fixed population size, n, and fixed mean number of friends per individual, k, neither the average path length $L(p_r)$ nor the average clustering coefficient $C(p_r)$ in the social network significantly affect the p-value. For example, the p-value is almost one in the Uniform model, meaning that, under the assumption of no homophily in the social network, an individual barely changes his or her decision regardless of whether the social network is a regular network (when $p_r = 0$ and the average path length and the average clustering coefficient are large), a small-world network (when $0.001 < p_r < 0.15$ and the average path lenght is small and the average clustering coefficient is large), or a random network (when $p_r = 1$ and the average path length and the average clustering coefficient are small). In contrast, based on the assumption of no homophily in the social network, the p-value is almost zero in the Zipf model, indicating that an individual is very likely to be influenced by friends and change his or her intrinsic choice because of the empathetic effect, regardless of the social network structure.

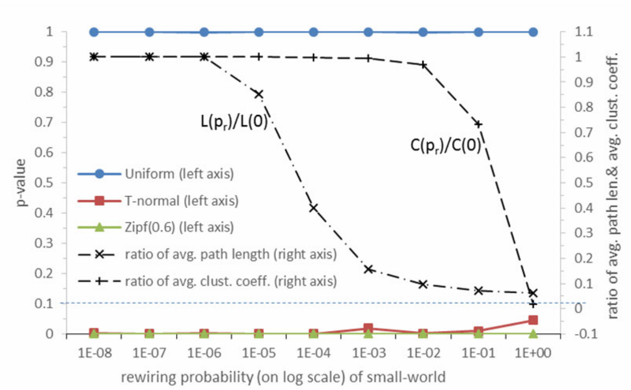


Figure 9. Correlation among p-value, ratio of average path length, ratio of average clustering coefficent and rewiring probability (on log scale) of small-world (n = 4000, k = 50, m = 10, s = 100, p_h = 0, w_o = 1800, w_f = 200, w_p = 0, θ = 0.6)

5.6 Effect of Strength of the Empathetic Preference

Figure 10 plots how the strength of the empathetic preference under the influence of friends is related to the p-value. Evidently, a larger strength makes the empathetic effect more likely to occur, increasing the probability that an individual changes his or her decision and the p-value is reduced. Figure 11 presents how the strength of the empathetic preference under the influence of the population of a social network is related to the p-value; the results are similar to those in Figure 10, especially with respect to the significant effects in the T-normal and Zipf models. As presented in Figure 11, when the strength exceeds two, the p-value equals almost zero, meaning that the distribution of people's final choices differs from that of their intrinsic choices.

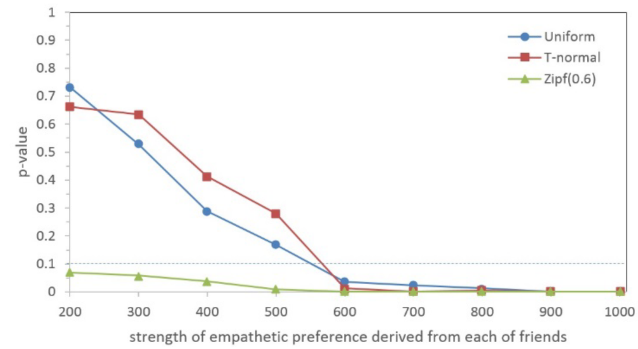


Figure 10. Correlation between P-value and strength of empathetic preference derived from each of friends (n = 4000, k = 50, m = 10, s = 100, p_h = 0.5, p_r = 0.15, w_o = 1800, w_p = 0, θ = 0.6)

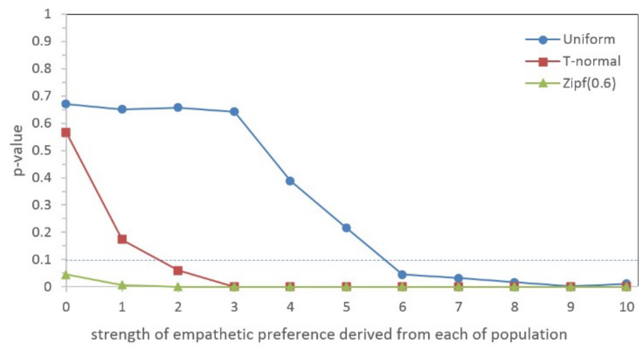


Figure 11. Correlation between P-value and strength of empathetic preference derived from each member of population (n = 4000, k = 50, m = 10, s = 100, p_h = 0.5, p_r = 0.15, w_o = 1800, w_f = 200, θ = 0.6)

5.7 Effect of Zipf Rank Exponent

Figure 12 plots the effect of the value of the Zipf rank exponent (θ), which is a parameter in the Zipf model, on the p-value. A larger θ value leads to a higher variance in the distribution of intrinsic choices for each of the available options. In this case, an individual is more likely to experience the empathetic effect when making his or her decision and choose the most chosen option. This phenomenon is referred to "the rich getting richer" or "preferential attachment" and results in a smaller p-value. In contrast, when the θ value is smaller, the intrinsic choices are more evenly distributed among alternatives and an individual is less likely to be influenced to change his or her decision. When θ = 0, the Zipf model approximates the Uniform model and the p-value is the largest.

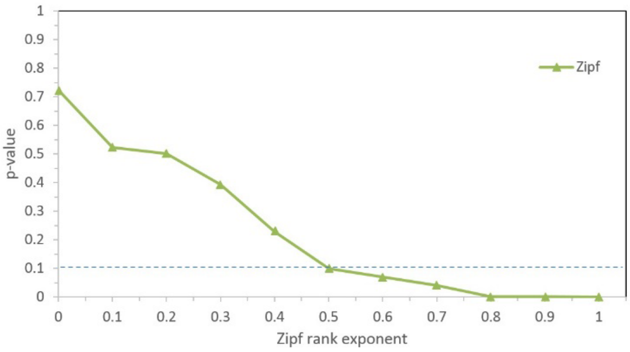


Figure 12. Correlation between *P*-value and Zipf rank exponent (n = 4000, k = 50, m = 10, s = 100, p_h = 0.5, p_r = 0.15, w_o = 1800, w_f = 200, w_p = 0)

6 Conclusions

In this paper, a model of social networks is constructed based on the small-world network and a simulation model of social networks with homophily is proposed for simulating empathetic decision-making behavior in individuals. Initially, the distribution of intrinsic choices among available options is defined using three probability models - uniform, truncated-normal (T-normal), and Zipf. Then, many simulation experiments are performed. Finally, the Chi-square goodness-of-fit test is conducted to test whether the distribution of people's final choices is the same as that of their intrinsic choices, revealing how experimental parameters influence the empathetic effect.

Experimental results reveal that population size does not significantly affect the empathetic effect in either the Uniform model or the T-normal model, and that the empathetic effect in the Zipf model becomes stronger as the population size increases, causing the distribution of people's intrinsic choices to be inconsistent with that of their final choices. The experiment also shows an individual with more friends is more likely to be influenced by friends and change his or her intrinsic choice.

The homophily parameter yields the same experimental results in the Uniform model, the T-normal model, and the Zipf model. In social networks with greater homophily, the empathetic effect is less likely to occur and decision-makers are more likely to make their final choices according to their intrinsic preferences.

The experimental results also indicate that decision-makers are less likely to experience the empathetic effect when they have more available options from which to choose.

The experiment that was also conducted on various social network structures reveals that the network structure does not significantly affect the experimental results. Regardless of the social network structure, the empathetic effect is weakest in the Uniform model, middling in the T-normal model and strongest in the Zipf model.

Stronger influence by friends and members of a social network populations constitutes a stronger empathetic effect, so decision-makers are more likely to be influenced by others into changing their decisions.

When the value of the Zipf rank exponent (θ) is greater, the difference between the number of people who choose different options as their intrinsic choice becomes more significant. Therefore, the individuals are more likely to experience the empathetic effect and select the most popular option.

Acknowledgments

The authors would like to thank the Ministry of Science and Technology of the Republic of China, Taiwan, for financially supporting this research under Contract No. MOST 104-2221-E-027-018-.

References

[1] D. Easley, J. Kleinberg, *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*, Cambridge University Press, 2010.

[2] M. McPherson, L. Smith-Lovin, J. M. Cook, Birds of a Feather: Homophily in Social Networks, *Annual Review of Sociology*, Vol. 27, pp. 415-444, August, 2001.

[3] S. Milgram, The Small World Problem, *Psychology Today*, Vol. 1, No. 1, pp. 60-67, May, 1967.

[4] R. Albert, H. Jeong, A. L. Barabási, Internet: Diameter of the World-Wide Web, *Nature*, Vol. 401, No. 6749, pp. 130-131, September, 1999.

[5] D. J. Watts, S. H. Strogatz, Collective Dynamics of "small-world" Networks, Nature, Vol. 393, No. 6684, pp. 440-442, June, 1998.

[6] J. Wohlgemuth, M. T. Matache, Small World Properties of Facebook Group Networks, *Complex Systems*, Vol. 23, No. 3, pp. 197-225, 2012.

[7] M. McPherson, L. Smith-Lovin, J. M. Cook, Birds of a Feather: Homophily in Social Networks, *Annual Review of Sociology*, Vol. 27, pp. 415-444, August, 2001.

[8] S. Aral, L. Muchnik, A. Sundararajan, Distinguishing Influence-based Contagion from Homophily-driven Diffusion in Dynamic Networks, *Proceedings of the National Academy of Sciences*, Vol. 106, No. 51, pp. 21544-21549, December, 2009.

[9] M. Yavaş, G. Yücel, Impact of Homophily on Diffusion Dynamics over Social Networks, *Social Science Computer Review*, Vol. 32, No. 3, pp. 354-372, March, 2014.

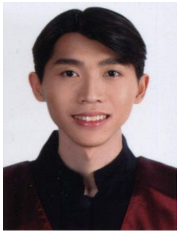
[10] N. M. van Gerwen, S. E. van Sleeuwen, Chosen or Meant to Be? A Study on the Difference between Opportunity-based and Choice-based Homogeneity on Gender and Socioeconomic Status in Facebook Networks, *Social Cosmos*, Vol. 4, No. 2, pp. 134-151, June, 2013.

[11] P. Pin, B. Rogers, Stochastic Network Formation and Homophily, Y. Bramoullé, A. Galeotti, B. Rogers (Eds.), *The*

Oxford Handbook of the Economics of Networks, Oxford, UK, 2016, pp. 303-357.

[12] A. Salehi-Abari, C. Boutilier, Empathetic Social Choice on Social Networks, *International Foundation for Autonomous Agents and Multiagent Systems*, Paris, France, 2014, pp. 693-700.

Biographies



Chien-Min Chen received his M.S. (2003) in Information Management from Nan Hua University (Chiayi, Taiwan). He is currently a Ph.D. student at National Taipei University of Technology (Taipei, Taiwan). His research interests include graph theory, reliability evaluation of distributed computing systems and social network analysis.

Min-Sheng Lin received his M.S. (1991) & Ph.D. (1994) in Computer Science & Information Engineering from National Chiao Tung University (Hsin Chu, Taiwan). He is currently a professor at National Taipei University of Technology (Taipei, Taiwan). His research interests include discrete mathematics, graph theory and social network analysis.