Summer 2014

Modeling profile-attribute disclosure in online social networks from a game theoretic perspective

Jundong Chen

MODELING PROFILE-ATTRIBUTE DISCLOSURE IN ONLINE SOCIAL NETWORKS FROM A GAME THEORETIC PERSPECTIVE

by

Jundong Chen, B.S., M.S., M.S., M.S.

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

COLLEGE OF ENGINEERING AND SCIENCE
LOUISIANA TECH UNIVERSITY

August 2014

LOUISIANA TECH UNIVERSITY

THE GRADUATE SCHOOL

June 23rd, 2014
Date

We hereby recommend that the dissertation prepared under our supervision by Jundong Chen

entitled

Modeling Profile-Attribute Disclosure in Online Social Networks from a Game Theoretic Perspective

be accepted in partial fulfillment of the requirements for the Degree of

Doctor of Philosophy

Vir Virander Phoha
Supervisor of Dissertation Research

Weizhong Dai
Head of Department

CAM
Department

Recommendation concurred in:

Weizhong Dai

Jinko Kanno

Matthias Ernst

Md Enamul Karim

R Selmic

Advisory Committee

Approved:

Director of Graduate Studies

Hisham Hegab

Dean of the College

Approved:

Sheryl S. Shoemaker

Dean of the Graduate School

MR

ABSTRACT

Privacy settings are a crucial part of any online social network as users are confronted with determining which and how many profile attributes to disclose. Revealing more attributes increases users' chances of finding friends and yet leaves users more vulnerable to dangers such as identity theft. In this dissertation, we consider the problem of finding the optimal strategy for the disclosure of user attributes in social networks from a game-theoretic perspective.

We model the privacy settings' dynamics of social networks with three game-theoretic approaches. In a *two-user* game, each user selects an ideal number of attributes to disclose to each other according to a utility function. We extend this model with a basic *evolutionary* game to observe how much of their profiles users are comfortable with revealing, and how this changes over time. We then consider a weighted evolutionary game to investigate the influence of attribute importance, benefit, risk and the network topology on the users' attribute disclosure behavior.

The two-user game results show how one user's privacy settings are influenced by the settings of another user. The basic evolutionary game results show that the higher the motivation to reveal attributes, the longer users take to stabilize their privacy settings. Results from the weighted evolutionary game show that: irrespective of risk, users are more likely to reveal their most important attributes than their least important attributes; when the users' range of influence is increased, the risk factor plays a smaller role in attribute disclosure; the network topology exhibits a considerable effect on the privacy in an environment with risk.

Motivation and risk are identified as important factors in determining how efficiently stability of privacy settings is achieved and what settings users will adopt given different parameters. Additionally, the privacy settings are affected by the network topology and the importance users attach to specific attributes. Our models indicate that users of social networks eventually adopt profile settings that provide the highest possible privacy if there is any risk, despite how high the motivation to reveal attributes is. The provided models and the gained results are particularly important to social network designers and providers because they enable us to understand the influence of different factors on users' privacy choices.

APPROVAL FOR SCHOLARLY DISSEMINATION

The author grants to the Prescott Memorial Library of Louisiana Tech University the right to reproduce, by appropriate methods, upon request, any or all portions of this Thesis. It is understood that "proper request" consists of the agreement, on the part of the requesting party, that said reproduction is for his personal use and that subsequent reproduction will not occur without written approval of the author of this Thesis. Further, any portions of the Thesis used in books, papers, and other works must be appropriately referenced to this Thesis.

Finally, the author of this Thesis reserves the right to publish freely, in the literature, at any time, any or all portions of this Thesis.

Author Jundong Chen

Date July 30, 2014

DEDICATION

This dissertation is dedicated to my beloved wife, Yanni Xu.

TABLE OF CONTENTS

ABSTRACT iii

DEDICATION vi

LIST OF TABLES ix

LIST OF FIGURES x

CHAPTER 1 INTRODUCTION 1

1.1 Overview 1

1.2 Dissertation Contributions 6

1.3 Definitions and Terminology 7

1.4 Organization of the Dissertation 9

CHAPTER 2 BACKGROUND AND RELATED WORK 10

CHAPTER 3 METHODS AND MODELS 14

3.1 Definitions and Strategies 14

3.1.1 Definitions 14

3.1.2 Strategies 18

3.1.3 Network Topologies 19

3.2 Models 21

3.2.1 The Two-user Attribute Disclosure Game 21

3.2.2 Basic Evolutionary Game 26

3.2.3 The Weighted Evolutionary Game 29

3.2.4 Working Case for Risk-free Scenario 33

CHAPTER 4 THEORETICAL ANALYSIS 38

CHAPTER 5 RESULTS AND DISCUSSION 42

5.1 Simulations Settings 42

5.1.1 Two-user Game 42

5.1.2 Basic Evolutionary Game 44

5.1.3 Weighted Evolutionary Game 45

5.2 Results 47

5.2.1 Two-user Game 47

5.2.2 Basic Evolutionary Game 50

5.2.3 Weighted Evolutionary Model 54

CHAPTER 6 CONCLUSIONS AND FUTURE WORK 67

BIBLIOGRAPHY 69

LIST OF TABLES

Table 3.1: Assumptions and Justifications for the Models 22

Table 3.2: The Process of Calculating Payoff Value and Choosing Mimic Object from the Candidate Neighbors. 35

Table 5.1: Parameter Values Used in the Models 43

Table 5.2: Payoff Matrix of the Two-user Game for Risk-included Scenario 43

Table 5.3: Payoff Matrix of the Two-user Game for Risk-free Scenario............... 44

Table 5.4: The Properties of Networks in the Simulation.................................... 46

LIST OF FIGURES

Figure 3.1: The figure shows a possible profile configuration for two users x and y, who disclose k_x and k_y attributes, respectively, from m possible attributes. The clear rectangles represent the disclosed attributes while the shaded rectangles represent withheld the attributes. 15

Figure 3.2: The figure illustrates the concept of equal value pairs given a profile configuration between two users x and y who share r pairs. An equal value pair is the occurrence of an identical valued attribute between two users. If the value of an attribute *Attr*$\#i$ is represented by v_i, then the figure shows that attributes *Attr*$\#1$...*Attr*$\#\eta$ have identical values for both users $(v_1...v_\eta)$ and therefore make up η equal value pairs. The attributes *Attr*$\#(\eta + 1)$...*Attr*$\#r$ compose the pairs that are not equal value pairs since they do not have identical values in both users. 16

Figure 3.3: The risk to a user's identity is dependent on whether that user can be easily distinguished from the rest of a network. User x is *hidden* if another user in the network exhibits characteristics identical to User x. Given users x and y disclose k_x and k_y attributes, respectively, User x can be "hidden" by User y if $k_x \leq k_y$ and the values exhibited by the k_x attributes are identical to the same attributes in User y's profile. The figure shows such a scenario in which attributes *Attr*$\#1$...*Attr*$\#k_x$ exhibit values $v_1...v_{k_x}$ that are identical to the first k_x attributes revealed by User y. In this case, User x is hidden by User y 17

Figure 3.4: The figure shows the two categories of users who can "hide" User x. The first category discloses the same number of attributes as User x, where all the revealed attributes are identical in value. The second category of users discloses more attributes in addition to the attributes disclosed by User x. The value of attribute *Attr*$\#i$ is denoted by v_i. Disclosed attributes are represented by clear squares, while the withheld attributes are represented by shaded squares. 18

Figure 3.5: The figure shows the probability of being friends from Equation 3.3 for different values of motivation α and equal value pairs η. Increasing the number of equal value pairs boosts the probability of being friends for all α values. This observation mimics the *homophily* principle more commonly known as "birds of a feather flock together," since increased equal value pairs indicates similarity between users. 24

Figure 3.6: In the basic evolutionary game, the population is divided into $m+1$ groups, which correspond to how many attributes are revealed. Given that every user has m attributes, their strategy options include revealing 1, 2, ..., m attributes or withholding all attributes (revealing 0 attributes), hence a total of $m + 1$ strategies. We calculate the positive and negative utilities of an arbitrarily selected User δ_k by comparing them with different groups of users. 27

Figure 3.7: The figure shows the implementation of selecting one of the neighbors as y^* based on the model of *balls into non-uniform bins*, where $C_x = y_1, y_2, ..., y_{|C_x|}$. The probability of selecting neighbor y_i is directly proportional to P^{t+1}_{x,y_i}. 33

Figure 3.8: (a) A sample network consisting of eight users, (b) each user has a profile with seven attributes with a weight vector, (c) initial sign flags for all eight users, (d) every user updates their strategy, and (e) the illustrated system converges. 37

Figure 5.1: The network topologies used in the simulations. The average node degree for each network is 4, and each network includes 100 nodes. .. 47

Figure 5.2: The Facebook friend networks used in the simulations. Network FB1 and FB2 are comprised of 151 and 502 nodes, respectively. 47

Figure 5.3: The correlation map shows the Nash equilibrium states for the two-player game with the colors corresponding to the probability of any user taking a specific strategy when Nash equilibrium is attained. Dark colors represent higher probabilities, and lighter colors represent lower probabilities. 49

Figure 5.4: Population dynamics for the basic evolutionary game for the risk-included scenario with different levels of motivation. 51

Figure 5.5: Population dynamics for the basic evolutionary game for the risk-free scenario with different levels of motivation. 52

Figure 5.6: Attribute dynamics for the weighted evolutionary game in random network, where the influential range of the utility function includes Friends. 57

Figure 5.7: Attribute dynamics for the weighted evolutionary game in random network, where the influential range of the utility function includes Friends and Friends-of-Friends. 58

Figure 5.8: Attribute dynamics for the weighted evolutionary game in small-world network, where the influential range of the utility function includes Friends. 59

Figure 5.9: Attribute dynamics for the weighted evolutionary game in small-world network, where the influential range of the utility function includes Friends and Friends-of-Friends. 60

Figure 5.10: Attribute dynamics for the weighted evolutionary game in scale-free network, where the influential range of the utility function includes Friends. 61

Figure 5.11: Attribute dynamics for the weighted evolutionary game in scale-free network, where the influential range of the utility function includes Friends and Friends-of-Friends. 62

Figure 5.12: Attribute dynamics for the weighted evolutionary game in FB1, where the influential range of the utility function includes Friends. ... 63

Figure 5.13: Attribute dynamics for the weighted evolutionary game in FB1, where the influential range of the utility function includes Friends and Friends-of-Friends. 64

Figure 5.14: Attribute dynamics for the weighted evolutionary game in FB2, where the influential range of the utility function includes Friends. ... 65

Figure 5.15: Attribute dynamics for the weighted evolutionary game in FB2, where the influential range of the utility function includes Friends and Friends-of-Friends. 66

CHAPTER 1

INTRODUCTION

1.1 Overview

Concerns regarding the privacy in social networks have received worldwide attention and led to frequent public debates [1, 2]. Social networks contain large amounts of information that can be used uniquely to identify their users as well as provide information on their habits, interests, and history [3]. On the positive side, the information enables the users to identify potential new "friends" and find old friends [4]. However, revealing information also makes it accessible to potential criminals, leaving the users vulnerable to dangers such as identity theft, sexual predators, stalkers, and inference by defrauders [19]. The risk to user privacy has caused so much concern that over 60% of social network users employ privacy increasing measures such as deleting friends and concealing profile attributes from other social network users [20]. The benefits and risks create a dilemma that every user of a social network faces: reveal more attributes to attract more friends, or reveal less attributes and become less vulnerable.

A considerable amount of research has been done in understanding online social networks and the factors that contribute towards their success. Online social networks are built on the concept of *self-disclosure* [21], which is positively affected by factors like *relationship-building* and *platform enjoyment.* In contrast, perceived privacy risk is a factor with a negative effect on self-disclosure [21]. The benefit of relationship-building is linked to the number of friends a user stands to gain by disclosing personal information. The link between number of potential friends and revealed information is based on the *homophily* principle more commonly expressed as "birds of a feather flock together" [22]. In the context of a social network, this principle translates to users with similar attributes being more likely to establish a friendship [22, 23]. On top of the similarity in attributes, the number of revealed attributes also positively affects relationship-building. Lampe *et al.* [24] find that the number of friends that a user has is exponentially related to the size of the set of attributes that the user reveals. This is because sharing more profile attributes allows more users to establish common ground that promotes interaction and encourages "friendship" [25]. However, *profile disclosing* increases the privacy risk to social network users [21]. Profile disclosing is defined as the amount of a user's profile that is visible to a third party [21, 26].

Therefore, each user in a social network weighs both the risks and benefits to determine how many profile attributes to reveal. Additionally, the privacy setting of one user affects the choice of privacy setting of another user. However, little work has been done to show how all these factors are linked together. Consequently, there is a need to model the interaction of users in a generic social network to understand how privacy risk and relationship-building both influence the level of self-disclosure exhibited in that network. Such a model would be invaluable in predicting the general preference of users when it comes to privacy in social networks.

Game theoretic models have been applied to online social networks before. Squicciarini *et al.* [27] present a general sum game involving a user and a server to explore the dynamics of user registration in social networks. The model and the results show that most users agree to provide their personal information during registration if the service provider promises to protect the users' privacy.

Evolutionary games have also been applied to social networks. Using the results from a survey, Squicciarini *et al.* [28] build an evolutionary game theoretical model aimed at optimizing the users' long-term utility. Additionally, by investigating the evolutionary game dynamics, they discover that social capital gained from self-disclosure influences a user's decisions more than the risk to that user's privacy.

The profile attribute privacy problem is similar to the stag-hunt game which exhibits both pure and mixed Nash equilibria [29]. The stag-hunt game is a two-player two-strategy game that captures the conflict between cooperation and safety involved in a situation where a hunter selects whether to hunt a stag or a hare without prior knowledge of another hunter's choice. This game reaps the maximum benefit to both players if both players select to hunt a stag and there is maximum risk to one player if the other player selects otherwise. This situation is similar to the privacy in social networks between two players because maximum benefit is accrued if both players cooperate and reveal all their attributes. Maximum risk occurs to a user when the other user reveals less attributes because this leaves the more revealed user vulnerable to identity inference. However, the profile attribute privacy problem is different from the stag-hunt game because the privacy problem involves multiple players, and multiple strategies (options) in their privacy.

Other works have also employed game theoretic models to capture the relation and coordination between different user properties in different networks in a variety of applications. The networks range from online video sharing social networks [30] to mobile ad-hoc networks [31], and anonymous social networks [32]. The modeled applications include sharing co-owned pictures in a social network [33] and stimulating cooperation in the network [30]. In most of these works [30, 34], a two party model is captured and used as a basis to create a model that captures the dynamics of the entire network. This is because the networks can be looked at as a collection of multiple two party interactions. We employ this same reasoning when designing models to capture the interaction of a user's privacy in a social network.

In this dissertation, we propose three game-theoretic models to study the dynamics of privacy settings between users in a social network. These models include a two-user model and two evolutionary game models and are built on a novel analytical definition of risk and motivation in a social network.

The *two-user game* models the interactions between two users in both risk-free and risk-included scenarios. We use this model to understand how the privacy choices of one user affect the privacy choices of another user. For example, given a network in which Alice and Bob are "friends" with an identical number of profile attributes, the two-user game investigates whether a strategy by Alice to withhold 30% of her attributes would make Bob withhold or reveal more of his attributes.

The *evolutionary game* is an extension of the two-user game to model the interactions of multiple users over time with the utility function of the evolutionary game derived from the utility function of the two user game. In the basic evolutionary game, all the users are allowed to change their strategy over time in order to maximize the benefit of friendship establishment and minimize the risk to their privacy. The users' choice of strategy at any point in time is dependent on the strategies currently employed by all the users in the network. Informally, given that Alice is part of a large social network, the evolutionary game investigates whether she would change her decision to withhold 30% of her attributes if she knew that 60% of the network's users were revealing all their attributes. The evolutionary game also investigates whether and how her new privacy strategy would affect the rest of the network users. This iterative process is repeated until the entire population reaches an equilibrium state. The equilibrium states as well as the dynamics of the network provide insights into understanding the privacy preferences of social network users.

The weighted evolutionary model also considers different types of networks and different types (weights) of attributes. This model investigates two concepts. Firstly, it investigates what influence, if any, the type of network has on the privacy strategy of the users of the network given the benefit of friendship enhancement. The network types considered include random networks, scale-free networks, and small-world networks to model different social network properties [35]. For example, given Alice is the popular girl in the social network and is a friend to everyone, this model investigates whether her strategy to withhold 30% of her attributes affects other users' strategies as much as Bob's decision, given he is less popular with only two friends in that network. Secondly, this model investigates whether the importance of the revealed and hidden attributes play a role in the decision. By weighing the attributes, this model considers the possibility that some attributes have a higher impact than others in either self-disclosure or privacy. This model investigates whether Alice revealing attributes such as her religion and sexual preferences would affect the network more than her revealing that she likes playing soccer and watching movies.

1.2 Dissertation Contributions

In this dissertation, we present the *Nash equilibria* [36] for the proposed two-user game model as well as the population dynamics for the evolutionary models. In our models, the Nash equilibrium refers to the optimal strategies taken by the users of the network. The strategies are optimal because the users cannot achieve a higher benefit by unilaterally changing their strategy. We also present the population dynamics for the evolutionary game showing the popularity of different strategies as different users change their privacy over time.

1. For the two-user game, we find that the pure strategy is for at least one of the players to disclose no attributes at all if there is an element of risk. Surprisingly, removing the risk element does not mean that all players will disclose all their attributes.

2. For the basic evolutionary game, we discover that the dominant strategy is to disclose no attributes if there is an element of risk. By dominant strategy, we refer to the strategy employed by most of the users in the social network. On the other hand, if the risk factor is ignored, the dominant strategy is to disclose all attributes. Revealing all but one attributes is also a common initial strategy in a risk-free network. Additionally, we find that networks where the risk factor is considered achieve equilibria faster than networks where risk was ignored.

Our results indicate that users will only be satisfied with the maximum privacy setting regardless of the motivation and benefits of less private settings as long as there is an element of risk in the social network.

3. Using the weighted evolutionary game model, we observe a tendency by users to reveal their most important attributes more than their less important attributes. By important attributes, we refer to those attributes which have a larger impact on the privacy as well as the social capital of a user. Additionally, users in random and scale-free networks are more likely to reveal their attributes than users in small-world networks. Interestingly, we find that the type of network topology has a limited effect on privacy settings of a social network in the risk-free case and yet have a considerable effect on the privacy in the risk-included scenario.

1.3 Definitions and Terminology

In this section we define the various terminologies that are central to the methodology used in this dissertation. Some of these terms are further described when they are first used in the dissertation.

Game: An interaction between rational, mutually aware players, where the payoffs of some players are influenced by the decisions of others [5].

Attribute: A field in a user's online social network profile. The importance of an attribute to a user is linked to the benefit gained from its revelation to other users in the network.

Benefit: In the two-user and basic evolutionary games, the benefit is captured by the expected number of potential friends that a user can make. In the weighted evolutionary game, the benefit is captured by the enhancement of a "friendship". We quantify this enhancement using the number and importance of attributes that a user shares with the neighbors.

Identity: The complete set of all profile attribute values of a user in a social network that differentiate that user from any other user in the network.

Motivation: A factor that captures the incentive for users to disclose profile attributes and affects benefit.

Privacy settings: A configuration of the social network users' profile information to enable or disable the visibility of certain profile attributes.

Risk (Privacy risk): The probability of a user's identity being inferred. It is inversely related to the number of the users who disclose the same attributes or additional attributes.

Strategy: A set of actions that players can follow. The strategy in two-user and basic evolutionary game models refers to how many attributes should be disclosed. In the weighted evolutionary game model, the strategy refers to which and how many attributes should be disclosed.

Utility: A quantity which represents the players' preference of a certain strategy. In our game model, utility includes benefit (positive utility) and risk (negative utility).

Random network: A network that is obtained by randomly sampling from a collection of networks which are constructed by the same amount of edges and vertices.

Small-world network: A mathematical graph where most nodes can be reached by other nodes in a small number of hops even though most of the nodes are not adjacent to each other.

Scale-free network: A network whose degree distribution follows the power law.

1.4 Organization of the Dissertation

In Chapter 2, we discuss the various past works which relate to privacy settings of online social networks and game theoretic models in this dissertation. We describe our game-theoretic models and specify the definition and strategies used in the models in Chapter 3. In Chapter 4, we provide theoretical analysis for the models. We then present the results and highlight the significance of our approach in Chapter 5 and conclude this dissertation with a discussion of our findings in Chapter 6.

CHAPTER 2

BACKGROUND AND RELATED WORK

A considerable amount of research has been done in understanding online social networks and the factors that contribute towards their success. Online social networks are built on the concept of *self-disclosure* [21], which is positively affected by factors like *relationship-building* and *platform enjoyment*. In contrast, perceived privacy risk is a factor with a negative effect on self-disclosure [21]. The benefit of relationship-building is linked to the number of friends a user stands to gain by disclosing personal information. The link between the number of potential friends and revealed information is based on the *homophily* principle more commonly expressed as "birds of a feather flock together" [22]. In the context of a social network, this principle translates to users with similar attributes being more likely to establish a friendship [22, 23].

On top of the similarity in attributes, the number of revealed attributes also positively affects relationship-building. Lampe *et al.* [24] find that the number of friends that a user has is exponentially related to the size of the set of attributes that the user reveals. This is because sharing more profile attributes allows more users to establish common ground that promotes interaction and encourages "friendship" [25]. However, *profile disclosing* increases the privacy risk to social network users [21].

Profile disclosing is defined as the amount of a user's profile that is visible to a third party [21, 26].

Therefore, each user in a social network weighs both the risks and benefits to determine how many profile attributes to reveal. Additionally, the privacy settings of one user affect the choice of privacy settings of another user. However, little work has been done to show how all these factors are linked together. Consequently, there is a need to model the interaction of users in a generic social network to understand how privacy risk and relationship-building both influence the level of self-disclosure exhibited in that network. Such a model would be invaluable in predicting the general preference of users when it comes to privacy in social networks.

Game theory is the analysis of situations involving conflicts of interest using mathematical models [14]. Each participant is referred to as a player, and each player has a set of possible strategies they can employ to achieve their goals. Each player's utility is jointly determined by the strategies chosen by all the players in the game. Game theory is a growing field that has been applied to many areas including various aspects of online social networks. These aspects range from modeling network formation [13], to community detection [15], and discovering influential nodes [16].

Game theoretic models have been applied to online social networks before. Using results from a survey, Squicciarini *et al.* [28] built an evolutionary game theoretic model aimed at optimizing the users' long-term utility. Additionally, by investigating the evolutionary game dynamics, they discovered that social capital gained from self-disclosure influences a user's decisions more than the risk to that user's privacy.

In [6, 7], we apply a weighted evolutionary game to model privacy settings of an online social network. This model captures the relative importance of profile attributes by assigning different weights to different attributes. The model also considers different types of network topologies. We discover that network connectivity and attribute importance have an effect on the disclosure of profile attributes.

In [8], in addition to the weighted evolutionary game, we investigate the one-to-one interplay in selecting the strategies on privacy settings by employing a two-user game. We also explore the influence of a motivation factor on the population dynamics by a basic evolutionary game.

The profile attribute privacy problem is similar to the *stag-hunt* game which exhibits both pure and mixed *Nash equilibria* [29]. The stag-hunt game is a two-player two-strategy game that captures the conflict between cooperation and safety involved in a situation where a hunter selects whether to hunt a stag or a hare without prior knowledge of another hunter's choice. This game reaps the highest benefit to both players if both of them decide to hunt a stag and there is a higher risk to one player if the other player selects otherwise. This situation is similar to the privacy in social networks between two players because the highest benefit is accrued if both players cooperate and reveal all their attributes. The highest risk occurs to a user when the other user reveals less attributes because this leaves the more revealed user vulnerable to identity inference. However, the profile attribute privacy problem is different from the stag-hunt game because the privacy problem involves multiple players, and multiple strategies (options) in their privacy.

Other works have also employed game theoretic models to capture the relation and coordination between different user properties in different networks in a variety of applications. The networks range from online video sharing social networks [30] to mobile ad-hoc networks [31], and anonymous social networks [32]. The modeled applications include sharing co-owned pictures in a social network [33] and stimulating cooperation in the network [30]. In most of these works [30, 34], a two party model is derived and used as a basis to create a model that captures the dynamics of the entire network. This is because the networks can be looked at as a collection of multiple two party interactions. We employ this same reasoning when designing models to capture the interaction of the user's privacy in a social network. We do not model the privacy settings of any specific online social network, but rather focus on a possible model for a generic online social network.

CHAPTER 3

METHODS AND MODELS

3.1 Definitions and Strategies

3.1.1 Definitions

In our model, the vector $A_x = (a_{x,1}, a_{x,2}, ..., a_{x,m})$ denotes the profile attributes in the social network, where $a_{x,i}$ is the i^{th} attribute of User x. An example of an attribute vector for a generic user (Alice) is given by $A_{Alice} = (Name, Gender, Age, Religion, ..., Hometown)$. For simplicity, we assume all the users have the same set of profile attributes. In a generic case, we refer to a specific attribute by *Attr#i.*

The value of the attributes is defined as a mapping $A_x \rightarrow V_x$, where $V_x = (v_{x,1}, v_{x,2}, ..., v_{x,m})$ is a vector of the values of the attributes of User x. We use $v_{x,i}$ to denote the value of i^{th} attribute of User x. For example, $V_{Alice} = (Alice, Female, 27, Christian, ..., Chicago)$.

For each User x, a vector $S_x = (s_{x,1}, s_{x,2}, ..., s_{x,m})$ denotes whether specific attributes are disclosed or revealed. If attribute $a_{x,i}$ is disclosed, then $s_{x,i} = 1$; otherwise, $s_{x,i} = 0$. For example, $S_{Alice} = (1, 1, 0, 0, ..., 1)$ means that Alice decides to reveal her name, gender, and hometown, but withholds her age and religion.

We capture the similarities between two users using *pairs.* Two users Alice and Bob are said to have a *pair* if they both reveal the same attribute, e.g. age. Moreover, if both users have the same value for that mutually revealed attribute (e.g. the age for both of them is 27), then the users are said to have an *equal value pair*. Formally, a 2-tuple $(a_{x,i}, a_{y,i})$ is called a pair if and only if $s_{x,i} = 1$ and $s_{y,i} = 1$. Additionally, if $v_{x,i} = v_{y,i}$, then the 2-tuple $(a_{x,i}, a_{y,i})$ is referred to as an equal value pair. We use random variable N_p to represent the number of pairs that two users share and random variable N_{ep} as the number of equal value pairs of two users.

Figure 3.1 shows a possible profile configuration for two users x and y. Out of the m attributes, User x reveals k_x attributes while User y reveals k_y attributes. Both users reveal attributes $Attr\#1, Attr\#2, ..., Attr\#r$, which contribute to r pairs. The r pairs are denoted by $(a_{x,1}, a_{y,1})$, $(a_{x,2}, a_{y,2})$, ..., $(a_{x,r}, a_{y,r})$.

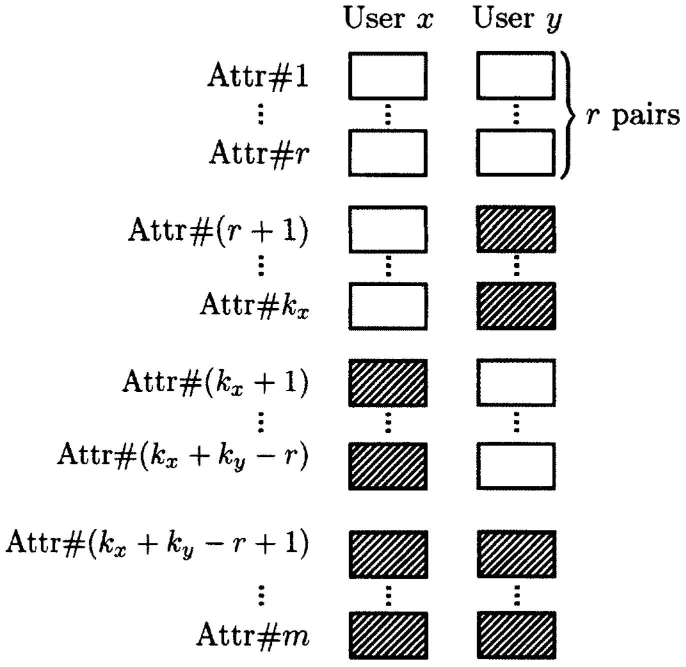


Figure 3.1: The figure shows a possible profile configuration for two users x and y, who disclose k_x and k_y attributes, respectively, from m possible attributes. The clear rectangles represent the disclosed attributes while the shaded rectangles represent withheld the attributes.

Figure 3.2 shows a similar profile configuration for the two users x and y with the attributes re-arranged such that the first r attributes are the r pairs. We consider that η of the r pairs are equal value pairs. We assume that the η equal value pairs are important in establishing common ground for building friendships [25].

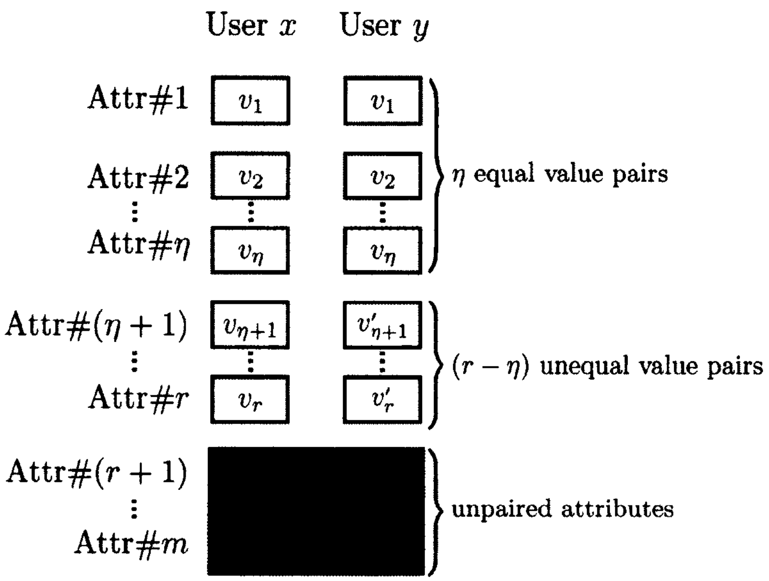


Figure 3.2: The figure illustrates the concept of equal value pairs given a profile configuration between two users x and y who share r pairs. An equal value pair is the occurrence of an identical valued attribute between two users. If the value of an attribute *Attr#i* is represented by v_i, then the figure shows that attributes *Attr#*1...*Attr#*η have identical values for both users $(v_1 ... v_\eta)$ and therefore make up η equal value pairs. The attributes *Attr#*$(\eta+1)$...*Attr#r* compose the pairs that are not equal value pairs since they do not have identical values in both users.

To capture the risk of identity inference, we introduce the concept of *hiding*. A user John is *hidden* by another user Jane if Jane is more distinguishable than John. For example, if $V_{JohnDoe} = (Doe, *, 34, *, ..., *)$ and $V_{JaneDoe} = (Doe, Female, 34, *, ..., Chicago)$, where '$*$' refers to withheld attributes, then Jane is more distinguishable than John and therefore John is hidden by Jane. This is because it takes less effort for a third party to infer the identity of Jane than John given the revealed profile attributes. Formally, given User x discloses k_x attributes and User y discloses k_y, where $k_x \leq k_y$, User x is *hidden* by User y if all the k_x attributes disclosed by User x have the same values in both User x and y (cf. Figure 3.3). The set D_x of attributes disclosed by User x is given by $D_x = \{a_{x,i} \mid s_{x,i} = 1, 1 \leq i \leq m\}$. User x is hidden by User y if and only if $D_x \subseteq D_y$ and $v_{x,i} = v_{y,i}$ for all $a_{x,i} \in D_x$.

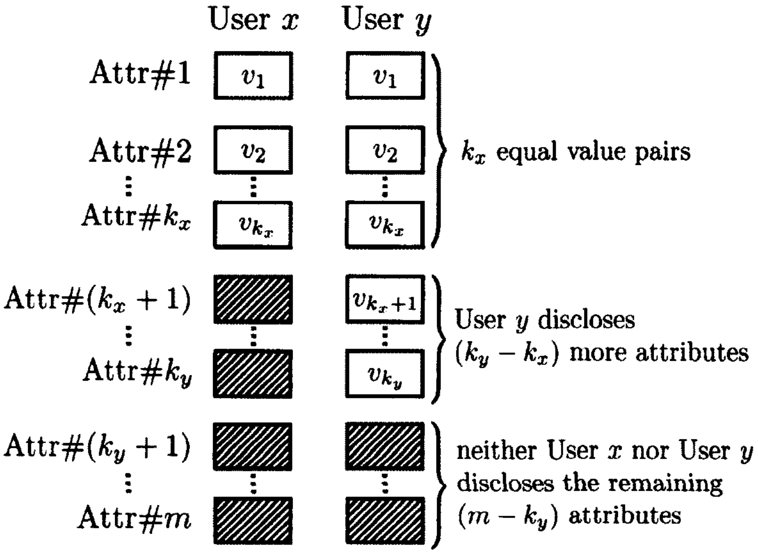


Figure 3.3: The risk to a user's identity is dependent on whether that user can be easily distinguished from the rest of a network. User x is *hidden* if another user in the network exhibits characteristics identical to User x. Given users x and y disclose k_x and k_y attributes, respectively, User x can be "hidden" by User y if $k_x \leq k_y$ and the values exhibited by the k_x attributes are identical to the same attributes in User y's profile. The figure shows such a scenario in which attributes *Attr#*1...*Attr#*k_x exhibit values $v_1 ... v_{k_x}$ that are identical to the first k_x attributes revealed by User y. In this case, User x is hidden by User y.

Therefore, User x can be hidden by two types of users (cf. Figure 3.4). One type of users consists of the users who disclose the same set of attributes, where corresponding attributes have identical values ($k_x = k_y = \eta$). The other type of users consists of those users who reveal extra attributes in addition to the k_x equal value pairs ($k_x = \eta < k_y$).

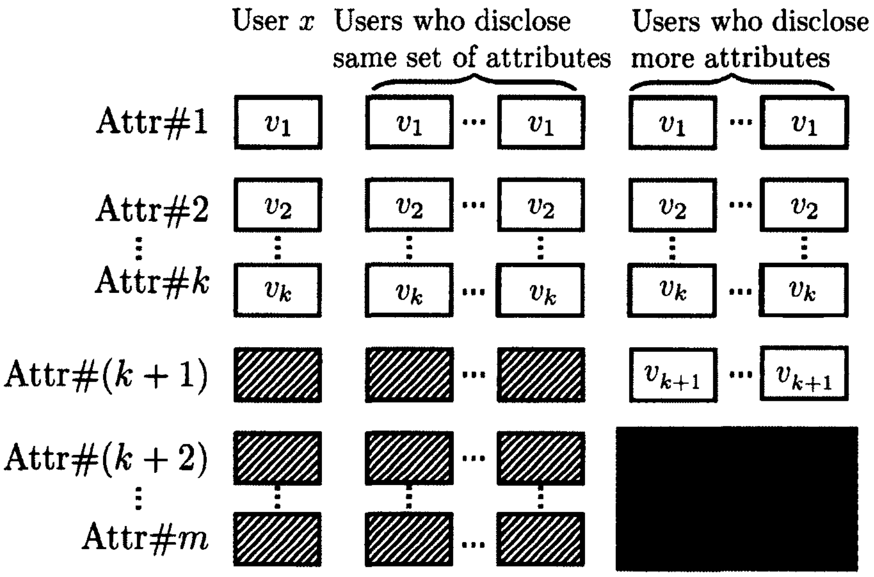


Figure 3.4: The figure shows the two categories of users who can "hide" User x. The first category discloses the same number of attributes as User x, where all the revealed attributes are identical in value. The second category of users discloses more attributes in addition to the attributes disclosed by User x. The value of attribute *Attr#i* is denoted by v_i. Disclosed attributes are represented by clear squares, while the withheld attributes are represented by shaded squares.

We define a social network as an undirected graph $G = (N, E)$ with node set N and edge set E, where the node set $N = \{1, 2, ..., n\}$ corresponds to n users in the network.

Additionally, we consider that the connectivity pattern of the network can follow the different network types described in the previous section. These networks include random, small-world, scale-free, and Facebook friend networks.

3.1.2 Strategies

The privacy settings of a typical social network consist of levels of visibility of different aspects such as profile attributes, activity logs, and friend lists to various types of users, e.g. friends, friends of friends, and public. In our model, we consider a single level of visibility, i.e. whether profile attributes are visible to any other user of the network.

In the two-user and basic evolutionary game, a user's strategy involves selecting how many attributes to reveal. Revealing more attributes increases the chance of having common attributes with other users which allows for friendship, while at the same time increases the risk of identity inference. Given m attributes, each user has $m + 1$ possible strategies which correspond to the number of attributes the user reveals (0, 1, ..., m). In the two-user game, we build an $(m + 1) \times (m + 1)$ payoff matrix made up of the payoff values for every possible strategy combination. The payoff values are evaluated from the positive and negative values associated with that strategy. In the basic evolutionary game, we classify the whole population into $m + 1$ groups depending on which strategy they adopt. Each group consists of users who have selected to reveal a given number of attributes.

In the weighted evolutionary game, the strategy involves selecting which and how many attributes to disclose.

3.1.3 Network Topologies

In this dissertation, the weighted evolutionary game considers three different types of network topologies, which include a random network, a small-world network, and a scale-free network.

A random network is a graph in which the occurrence of connection between nodes follows a probability distribution [37]. A random graph can be used for modeling social networks when the node degrees follow an arbitrary probability distribution. The Erdös-Rényi (ER) [39] model is considered to generate the random networks. The probabilities that edges occurs between any two nodes are equal and independent. Given the probability of an edge occurrence is p and there are n nodes, the average node degree k is approximately equal to $n \cdot p$.

In a small-world network, most of the nodes are not directly connected to each other, but most nodes can be reached by every other node within a relatively small number of hops. Online social networks have been shown to exhibit small-world properties and can be produced using a Watts-Strogatz model in two steps [40]. In the first step, a regular ring lattice of n nodes is created and each node is connected with $\frac{k}{2}$ on each side making the average node degree k. In the second step, the edges are rewired with probability β to create the "shortcuts" that transform the regular network to a small-world network [40].

A scale-free network is a network where the node degree distribution follows a power-law distribution, i.e. the number of nodes decreases exponentially as the node degree increases. The scale-free network is created using the preferential attachment mechanism, which means a node with a higher degree is more likely to attract new connections compared to a node with a lower degree [41].

3.2 Models

We propose three game-theoretic models. One model is a *two-user game*, which captures interactions between two users while setting up their privacy. This is extended to a *basic evolutionary game* to capture the dynamics of the privacy preference of multiple users in a large-scale social network. This model is then extended to a *weighted evolutionary game* to investigate the influence of attribute weight and network topology on the privacy of users in a social network.

There are many online social networks available today with a variety of privacy designs [9, 10]. Therefore, we model a generic social network with characteristics exhibited by some of the social networks. For example, in our models, every user has a profile made up of profile attributes, where each user is tasked with selecting how many and which attributes to reveal to other users. In the two-user and basic evolutionary games, the revelation is to all other users in the network, whereas in the weighted evolutionary game, the revelation is only to the user's friends. However, our models do not consider categories of friends with different levels of privacy which is a characteristic of some social networks. The assumptions used to construct the models and their justifications are provided in Table 3.1.

3.2.1 The Two-user Attribute Disclosure Game

In this model, we consider a two-user game between User x and User y to understand the basic interaction in complex networks such as online social networks. We use a utility function to capture the incentives of players [36].

Table 3.1: Assumptions and Justifications for the Models

Assumptions	**Justifications**
We define risk as the potential for identify inference.	Identity inference is an important path for privacy information leakage. Defining risk as the potential for identify inference allows us to build the relationship between profile attribute disclosure and privacy risk, even though there are other types of privacy risk issues.
We define the positive utility from information revelation as the number of friendships made in the two-user game and the basic evolutionary game.	With an increase in the number of friendships, the users can benefit more from communication with others, and sharing more information with others.
We assume equal importance for each profile attribute in the two-user and basic evolutionary games.	Many social network user profiles consist of similar attributes [11, 12]. In such a profile, one attribute is not necessarily more risky or important than another attribute. However, we also consider dissimilar attributes in the weighted evolutionary game model.
Users with more common attributes are more likely to be friends.	This assumption is based on research which shows that the *homophily* principle is exhibited in social networks [22, 23].
In the two-user and basic evolutionary games, we assume the probability of two users with η equal value pairs being friends is given by Equation 3.3.	Lampe *et al.* [24] find that the number of friends that a user has is exponentially related to the size of the set of attributes that user reveals.
In the two-user and basic evolutionary games, each attribute has the same number of possible values. In the weighted evolutionary game, all users in the network attach the same importance to any given attribute, e.g. all users will consider their address attribute to be more important than their religion attribute.	These assumptions allow us to investigate the influence of global network properties while simultaneously comparing local properties such as profile attributes and their importance to users on a common ground.

Positive Utility

The positive utility of revealing more attributes is the increased chance of establishing common ground with other users and thereby potentially obtaining more new friends. Given users x and y disclosing k_x and k_y attributes, respectively, the probability of having exactly r pairs is given by

$$P_{k_x,k_y}(N_p = r) = \frac{\binom{m}{k_x} \cdot \binom{k_x}{r} \cdot \binom{m-k_x}{k_y-r}}{\sum\limits_{i=\max\{k_x+k_y-m,0\}}^{\min\{k_x,k_y\}} \binom{m}{k_x} \cdot \binom{k_x}{i} \cdot \binom{m-k_x}{k_y-i}} \tag{3.1}$$

where m is the total number of attributes. We use random variables N_p and N_{ep} to denote the number of pairs and the number of equal value pairs, respectively. The proof for Equation 3.1 is provided in Section 4 (Theorem 4.1).

Given r pairs, we can calculate the probability of getting η equal value pairs, using

$$P(N_{ep} = \eta \mid N_p = r) = \frac{(L-1)^{r-\eta}}{L^r} \cdot \binom{r}{\eta} \tag{3.2}$$

where L is the number of values that an attribute can have. The proof for Equation 3.2 is provided in Section 4 (Theorem 4.2).

Given the number of friends that a user has is exponentially related to the set of attributes [24], we assume the probability of two users with η equal value pairs being friends is given by

$$P(F \mid N_{ep} = \eta) = \frac{e^{\alpha\eta}}{e^{\alpha m} + \epsilon} \tag{3.3}$$

where $\epsilon > 0$ and α indicates motivation. Dividing $e^{\alpha\eta}$ by $e^{\alpha m} + \epsilon$ guarantees that $P(F \mid N_{ep} = \eta)$ is between 0 and 1.

Additionally, we select an exponential style function because it mitigates the adverse effect brought by users with a lower number of equal value pairs. Users with a small set of equal value pairs are numerous but have little impact in building friendships.

The motivation α captures the increase in the likelihood of being friends with an increase in equal value pairs. We consider three different values for α. Figure 3.5 shows the likelihood of being friends evaluated from Equation 3.3 for $\alpha \in \{0.2, 0.9, 1.5\}$. Lower values of α have slower change but higher initial values.

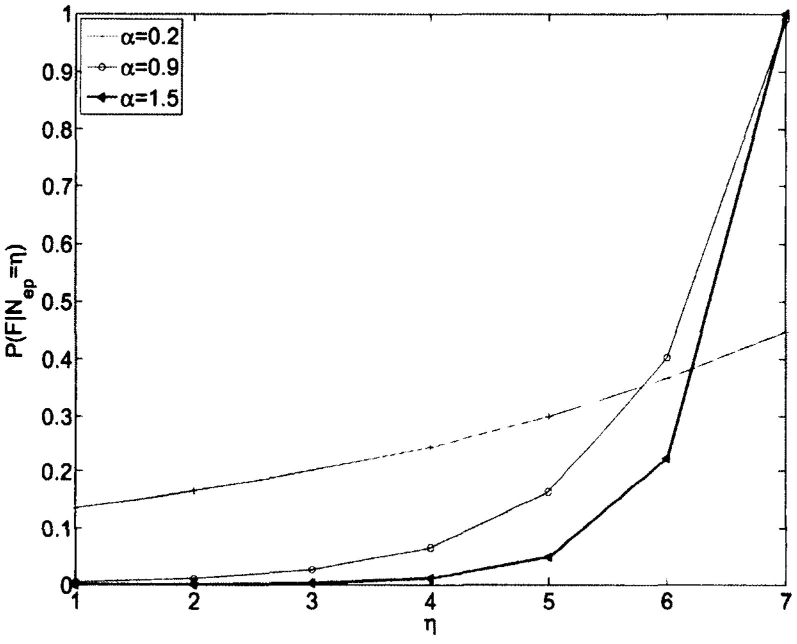


Figure 3.5: The figure shows the probability of being friends from Equation 3.3 for different values of motivation α and equal value pairs η. Increasing the number of equal value pairs boosts the probability of being friends for all α values. This observation mimics the *homophily* principle more commonly known as "birds of a feather flock together," since increased equal value pairs indicates similarity between users.

Given users x and y disclose k_x and k_y attributes, respectively, the probability $P_{k_x,k_y}(F)$ of them being friends is therefore given by combining two probabilities: (1) the probability of them being friends if they have η equal value pairs and (2) the probability of them having the η equal value pairs in the first place. $P_{k_x,k_y}(F)$ is evaluated using

$$P_{k_x,k_y}(F) = \sum_{r=\max\{k_x+k_y-m,0\}}^{\min\{k_x,k_y\}} P_{k_x,k_y}(N_p = r) \cdot \sum_{\eta=0}^{r} P(F \mid N_{ep} = \eta) \cdot P(N_{ep} = \eta \mid N_p = r), \tag{3.4}$$

where $P_{k_x,k_y}(N_p = r)$ is the probability of having exactly r pairs, $P(F \mid N_{ep} = \eta)$ is the probability of being friends given η equal value pairs, and $P(N_{ep} = \eta \mid N_p = r)$ is the probability of having η equal value pairs given r pairs.

Negative Utility

The negative utility of revealing more attributes is the increased risk incurred by the user. In our models, risk is equivalent to the chance of a user's identity being inferred from their disclosed attributes. A user's identity can be compared to a set of attributes that uniquely differentiate a user from a large group of users. Therefore, risk is inversely proportional to the number of users among whom a user can be hidden. This is because the higher the number of users with identical disclosed attributes, the less the probability of inferring a specific user's identity or preference. The users with unique sets of disclosed attributes have the highest risk in the population.

In the two-user game, User x discloses k_x attributes and User y discloses k_y attributes. If $k_x > k_y$, then there is no chance that User x is hidden by User y. The reason is that User x will always be more distinguishable than y regardless of which k_y attributes are selected by User y. If $k_x = k_y$ and all k_x disclosed attributes are equal value pairs, then User x is hidden by User y and they reduce each other's risk. This is because User x cannot be distinguished from y if all the attributes they reveal are identical. However, if $k_x < k_y$, and all k_x attributes disclosed by User x are all equal value pairs, then User x is hidden by User y, and the risk of User x is reduced. Hence, we get the formula for negative utility as follows:

$$R_{k_x,k_y} = \begin{cases} \frac{1}{1+\frac{\binom{m-k_x}{k_y-k_x}}{\binom{m}{k_y}}\frac{1}{L^{k_x}}} & \text{if } k_x \leq k_y, \\ 1 & \text{if } k_x > k_y, \end{cases} \tag{3.5}$$

where m is the total number of attributes. The proof for Equation 3.5 is provided in Section 4 (Theorem 4.3).

Combining Positive Utility and Negative Utility

When users x and y exist in the same social network, they have a probability of becoming friends with $P_{k_x,k_y}(F)$ and are also under the risk of identity inference R_{k_x,k_y}. We use the ratio of $P_{k_x,k_y}(F)$ to R_{k_x,k_y} to obtain an appropriate utility function

$$u_x = \frac{P_{k_x,k_y}(F)}{R_{k_x,k_y}}. \tag{3.6}$$

In Equation 3.6, User x discloses k_x attributes, while User y discloses k_y attributes.

3.2.2 Basic Evolutionary Game

A basic evolutionary game is employed to analyze the dynamics of privacy among multiple users in online social networks. Given m attributes, we divide the population into $m+1$ groups which consist of users who disclose the same number of attributes. Figure 3.6 shows an arbitrarily selected User δ_k who discloses k attributes and belongs to group k of n_k users.

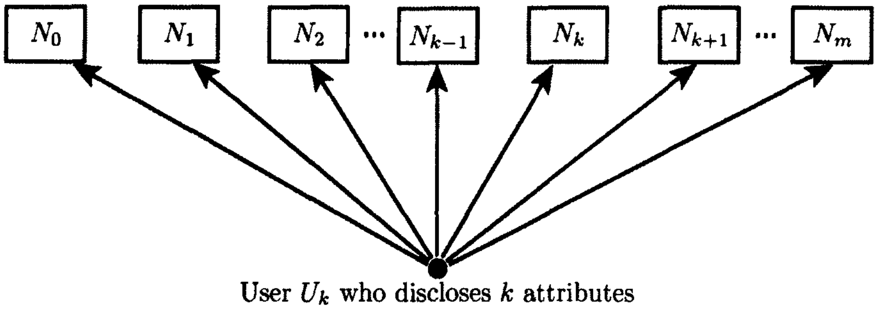


Figure 3.6: In the basic evolutionary game, the population is divided into $m + 1$ groups, which correspond to how many attributes are revealed. Given that every user has m attributes, their strategy options include revealing 1, 2, ..., m attributes or withholding all attributes (revealing 0 attributes), hence a total of $m + 1$ strategies. We calculate the positive and negative utilities of an arbitrarily selected User δ_k by comparing them with different groups of users.

Replicator Equation

Replicator dynamics are used to provide the population dynamics for each proportion [42]:

$$\dot{\theta_k} = \theta_k[f_k - \phi] \tag{3.7}$$

where θ_k is the proportion of all users who disclose k attributes. The parameter $\dot{\theta_k}$ is the differentiation of θ_k over time, where the time unit is the iteration step in the process of solving Equation 3.7. The value of θ_k is given by $\frac{n_k}{n}$, where n_k is the number of users disclosing k attributes and n is the total number of users. The fitness of type k is denoted by f_k which is defined later on in Equation 3.11, and ϕ is the *average population fitness* which is given by Equation 3.8.

Users who disclose k attributes are referred to as type k. The *fitness* of the basic evolutionary game is comparable to the utility function of the two-user game, and is also comprised of positive utility and negative utility. Positive utility is the expected number of friends that a user can make by disclosing k attributes while negative utility is the risk of inference from disclosing k attributes. Similar to the two-user game, the risk factor for a certain user is inversely proportional to the number of users that can hide that user.

The average population fitness ϕ is given by:

$$\phi = \sum_{k=0}^{m} \theta_k f_k. \tag{3.8}$$

Positive Utility

The positive utility is an extension of the two-user game's positive utility in Equation 3.4. In a large social network, the expected number of friends N_F that User δ_k can make is:

$$\begin{aligned} E_k[N_F] &= \sum_{k'=0}^{m} n_{k'} P_{k,k'}(F) \\ &= \sum_{k'=0}^{m} n_{k'} \sum_{r=\max\{k+k'-m,0\}}^{\min\{k,k'\}} P_{k,k'}(N_p = r) \\ &\cdot \sum_{\eta=0}^{r} P(F \mid N_{ep} = \eta) \cdot P(N_{ep} = \eta \mid N_p = r) \end{aligned} \tag{3.9}$$

where $n_{k'}$ is the number of users of type k'. Equation 3.9 is derived by summing Equation 3.4 for all possible k' values and respective $n_{k'}$.

Negative Utility

In the evolutionary game, the negative utility is calculated using

$$R_k = \frac{1}{\frac{1}{\binom{m}{k}} \cdot \frac{1}{L^k} \cdot (n_k - 1) + \sum_{i=1}^{m-k} \frac{\binom{m-k}{i}}{\binom{m}{k+i}} \cdot \frac{1}{L^k} \cdot n_{k+i}}. \tag{3.10}$$

The term $\frac{1}{\binom{m}{k}} \cdot \frac{1}{L^k} \cdot (n_k - 1)$ is the number of users of type k who can hide User δ_k. The term $\sum_{i=1}^{m-k} \frac{\binom{m-k}{i}}{\binom{m}{k+i}} \cdot \frac{1}{L^k} \cdot n_{k+i}$ is the number of users who can hide User δ_k, and are from the groups that disclose strictly more attributes.

Combining Positive Utility and Negative Utility

Similar to Equation 3.6, we use the ratio of positive utility to negative utility to define the fitness of type k using

$$f_k = \frac{E_k[N_F]}{R_k}. \tag{3.11}$$

All the users of the same type have the same fitness value.

3.2.3 The Weighted Evolutionary Game

We extend the previous model by considering a weighted evolutionary game. This model considers that users attach different importances to different attributes. This is captured by assigning weights to each attribute. Additionally, the topology of the network is considered. In this model, the positive utility of a user is affected by the number and type of attributes that a user shares with the neighbors.

We consider that the benefits and risks are affected by the users at two different levels of social closeness. The first level only includes User x's friends, and the second level also includes User x's *friends-of-friends*. We adopt *influential range* (IR) to represent which level of users contribute to User x's benefit and/or risk.

$$B_x(IR) = \begin{cases} \{F\}, & IR = 1, \\ \{F\} \cup \{FoF\}, & IR = 2, \end{cases} \tag{3.12}$$

where IR denotes influential range, F represents friend, and FoF stands for friend-of-friend. Therefore, $B_x(1)$ is the set of all the friends. $B_x(2)$ includes not just friends, but also friends-of-friends.

In our game, the utility is a combination of benefits (positive utility) and risks (negative utility). A user's positive utility is related to the amount and type of attributes that the user shares with other users in their influential range. The set of users who contribute to User x's positive utility is denoted by $B_x(IR)$.

Conversely, the risk is the probability of a user's identity being inferred. This probability is measured by the reciprocal of the number of the users who disclose the same or additional attributes, i.e. how many users in the influential range can hide that user. The set $B_x^h(IR)$ consists of users in the influential range who disclose the same attributes as x or extra attributes in addition to those disclosed by User x, and can possibly hide User x. The set $B_x^h(IR)$ determines how much risk a user is exposed to.

The combined utility (payoff) function is obtained by using

$$u_x = w_P \cdot \sum_{y \in B_x(IR)} (S_x \wedge S_y) \times W^T - w_N \cdot \frac{1}{|B_x^h(IR)|}, \tag{3.13}$$

where w_P and w_N are the weight coefficients for the positive utility $\sum_{y \in B_x(IR)} (S_x \wedge S_y)$ $\times W^T$ and negative utility $\frac{1}{|B_x^h|}$, respectively[1].

We define the benefit-to-risk ratio (BRR) as $w_P : w_N$, which is the ratio of the coefficient for positive utility to the coefficient for negative utility.

[1] Unless otherwise stated, we use notation $\wedge$ to represent logic AND. Notation W^T refers to the transpose of vector W.

Our model is iterative and synchronous. First, each user in the network is assigned a random initial attribute sign flag vector. In every iteration, each user compiles a set of candidate neighbors whose privacy settings they may mimic. This set consists of the neighbors who derive a higher utility from their privacy settings than the user derives from his/her own settings. Based on the neighbors' utilities, each user decides whether to change or maintain their strategy. A user is likely to change his/her strategy if his/her neighbors derive a higher utility from their own strategies than the user derives from his/her own. If a user decides to change his/her strategy, one of the candidate neighbors is then selected as the object to mimic. The mimicking process involves a user changing one digit of his sign flag to the corresponding digit of the candidate neighbor's sign flag. This is analogous to a user Alice deciding to reveal her location attribute after seeing that her friend Bob, who has a higher utility, has a revealed location attribute. At the end of each iteration, all the users update their strategies synchronously. The procedure keeps running iteratively until there are no users who change their sign flags between two consecutive iterations. When this condition has been met, the model is said to achieve convergence.

Formally, users follow the *replicator rule* to update their strategies between two successive time steps [43]. Each node makes a decision to maintain or change its current strategy based on the utilities exhibited by its neighbors. Given u_x^t and u_y^t are the utilities of User x and User y respectively at time t, the probability of User x (at time $t + 1$) adopting the strategy of User y (at time t) is given by:

$$P_{x,y}^{t+1} = \begin{cases} \frac{u_y^t - u_x^t}{d_{max}}, & u_y^t > u_x^t, \\ 0, & u_y^t \leq u_x^t. \end{cases} \tag{3.14}$$

We use the largest difference d_{max} in payoff between any two users in the network to guarantee that $P_{x,y}^{t+1} \in [0,1]$. Equation 3.14 implies that the probability of User x following the strategy of a neighbor (User y) is proportional to the payoff difference between users x and y, when User y's utility is higher than User x's. This probability value is evaluated for all members of the candidate neighbor set C_x.

Each user's decision to maintain or change his/her strategy depends on $P_{x,y}^{t+1}$ values for the entire candidate neighbor set C_x. The probability of User x maintaining its original strategy, as derived from [43], is given by:

$$\overline{Q_x^{t+1}} = \prod_{y \in C_x} (1 - P_{x,y}^{t+1}) \tag{3.15}$$

Conversely, the probability of User x changing its strategy between t and $t+1$ is given by

$$Q_x^{t+1} = 1 - \prod_{y \in C_x} (1 - P_{x,y}^{t+1}). \tag{3.16}$$

After evaluating all probabilities and deciding to change his/her strategy, each user selects the neighbor to mimic in the update process. A higher $P_{x,y}^{t+1}$ value for candidate y translates to a higher probability of being selected as the mimic object y^*. The implementation of selecting y^* is based on a mathematical model called *balls into non-uniform bins* [18], in which the probability[2] $P(y_i)$ of a ball falling into a certain bin is proportional to the size of the bin. In Figure 3.7, the size of the each

[2] In this dissertation, we use y to refer to a general user, and we use y_i to refer to a specific user.

bin is exactly equal to $P^{t+1}_{x,y}/\Delta$, where $\Delta = \sum_{y \in C_x} P^{t+1}_{x,y}$. In total, there are $|C_x|$ bins. Therefore, the probability of the ball falling into ith bin is given by:

$$P(y_i) = P^{t+1}_{x,y_i}/\Delta \tag{3.17}$$

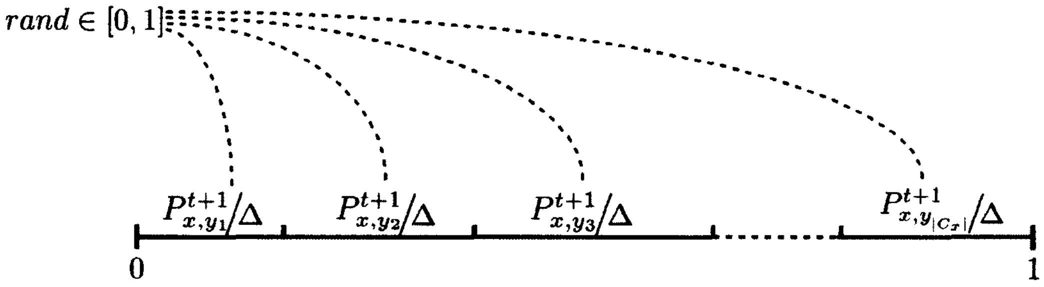


Figure 3.7: The figure shows the implementation of selecting one of the neighbors as y^* based on the model of *balls into non-uniform bins*, where $C_x = y_1, y_2, ..., y_{|C_x|}$. The probability of selecting neighbor y_i is directly proportional to P^{t+1}_{x,y_i}.

After the mimic object is determined, the specific attribute to mimic is randomly selected from the attributes with different sign values.

The algorithm for updating the attribute sign flag is shown in Algorithm 1.

3.2.4 Working Case for Risk-free Scenario

In this subsection, we describe a working case of a risk-free scenario of our model, in which the influential range is restricted to a user's friends (neighbors). Figure 3.8a shows the topology structure of the network in this example, which consists of 8 users, whose profile attributes and associated weights are shown in Figure 3.8b. The profile attributes include (Name, Gender, Age,..., Hometown) with weight vector $(w_1, w_2, w_3, ..., w_7) = (0.02, 0.06, 0.10, 0.14, 0.18, 0.22, 0.28)$. Figure 3.8c shows the initial sign flags for all 8 users. For example, User 5 has a sign flag $S_5 = (1100110)$, which means that only his/her name, gender, education, and occupation are revealed.

In the next few paragraphs, we show how User 5 may change his/her strategy in our model.

Algorithm 1: Algorithm for updating profile attribute sign flag

```
Input: Initial sign flag $iSF$
Output: Final sign flag $fSF$
Assign $iSF$ for each node;
do
    for each node do
        Find the set of candidate neighbors $C_x$;
        Evaluate $P_{x,y}^{t+1}$ for all members of candidate set;
        Evaluate probability of changing strategy $Q_x^{t+1}$;
        Generate a random number $rand \in [0, 1]$;
        if $rand < Q_x^{t+1}$ then
            /* Decision is made to change strategy */
            Select neighbor $y^*$ from $C_x$;
            /* Neighbor is selected using balls into non-uniform bins */
            Change single bit from $SF_x$ to mimic $SF_{y^*}$;
        end
    end
    All nodes update sign flags synchronously;
while any node changes sign flag;
return $fSF$
```

In the first step, every user calculates their utilities from Equation 3.13. This involves a comparison of the users' revealed attributes with each neighbor. User 5 has two neighbors: User 1 and User 2 with initial sign flags $S_1 = (1000110)$ and $S_2 = (0110011)$, respectively. The attribute pairs between any two users are obtained by using bit-wise AND operation between the users' sign flag vectors. The bit-wise AND operation between S_1 and S_5 is (1000110), which means that both User 1 and User 5 disclosed attributes 1, 5, and 6. The summation of the weights of attribute pairs (Equation 3.13) is therefore given by $w1 + w5 + w6$, which evaluates to 0.42. Similarly, the summation of the weights of attribute pairs between S_2 and S_5 is 0.28.

The positive utility for any user is obtained by summing the weighted pair sums for all his/her neighbors. In this case, the positive utility for User 5 is the sum of the weighted attribute pairs between User 5 and both User 1 and User 2. This evaluates to $0.42 + 0.28 = 0.70$. In a similar fashion, the utilities are evaluated for all the network users. Table 3.2 shows the positive utilities for Users 5, 1, and 2.

Table 3.2: The Process of Calculating Payoff Value and Choosing Mimic Object from the Candidate Neighbors.

User	**Neighbor**	**AND result**	**Weighted result**	**Positive utility**	$P_{x,y}^{t+1}$
User 5	User 1	1000110	$w_1 + w_5 + w_6 = 0.42$	0.70	N/A
	User 2	0100010	$w_2 + w_6 = 0.28$		
User 1	User 2	0000010	$w_6 = 0.22$	1.26	0.41
	User 3	1000100	$w_1 + w_5 = 0.2$		
	User 4	0000000	0		
	User 5	1000110	$w_1 + w_5 + w_6 = 0.42$		
	User 8	1000110	$w_1 + w_5 + w_6 = 0.42$		
User 2	User 1	0000010	$w_6 = 0.22$	1.38	0.49
	User 5	0100010	$w_2 + w_6 = 0.28$		
	User 6	0100011	$w_2 + w_6 + w_7 = 0.56$		
	User 7	0010000	$w_3 = 0.10$		
	User 8	0000010	$w_6 = 0.22$		

In the second step, each user evaluates the probability $P_{x,y}^{t+1}$ of mimicking his/her neighbors according to Equation 3.14. The maximum range between the utility values for the network nodes $d_{\max}$ is found to be 1.38. User 5 only has to consider User 1 and User 2 when evaluating these probability values. $P^1_{5,1}$ evaluates to 0.41 while $P^1_{5,2}$ evaluates to 0.49.

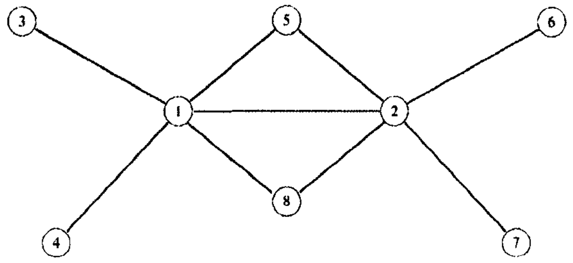
In the third step, each user decides whether to change or maintain his/her strategy by using Equation 3.16 which utilizes the probabilities evaluated in the step above. For User 5, Q^1_5 evaluates to 0.6991. If a randomly selected number in the range $[0, 1]$ is less than Q^1_5, then User 5 decides to change his/her strategy. Otherwise, User 5 maintains his/her strategy. In our case, User 5 decides to change his/her strategy.

In the fourth step, users who decided to change their strategies select a candidate neighbor to mimic. Candidate neighbors should exhibit higher utility values than the user itself. The probability of User x selecting a specific neighbor y is directly proportional to $P^{t+1}_{x,y}$ for that neighbor. Since Users 1 and 2 both have higher utilities than User 5, they are both viable candidates for User 5 to mimic. After normalizing $P^1_{5,1}$ and $P^1_{5,2}$, the bin sizes for User 1 and User 2 are 0.46 and 0.54, respectively (cf. Equation 3.17 and Figure 3.7). In our case, User 5 selects User 2 as the mimic object.

In the fifth step, each user who decided to change their strategy selects which attribute to reveal or withhold to resemble their mimic object. Comparing User 5's and User 2's sign flags reveals that they differ in four positions, i.e. 1, 3, 5, and 7. User 5 can mimic User 2 in one of the following ways: revealing attribute 3, revealing attribute 7, withholding attribute 1, or withholding attribute 5. In our case, User 5 decides to reveal attribute 7.

All five steps are repeated in each iteration until no single user changes his/her strategy between two successive iterations. The system is then said to have converged.

Figure 3.8d shows the sign flags for all eight users after a single iteration. Figure 3.8e shows the sign flags for the whole network after convergence. In this simulation, convergence is achieved after 11 iterations.



(a)

w_1	w_2	w_3	w_4	w_5	w_6	w_7
(0.02)	(0.06)	(0.10)	(0.14)	(0.18)	(0.22)	(0.28)
Name	Gender	Age	Religion	Education	Occupation	Hometown

(b)

Node	w_1	w_2	w_3	w_4	w_5	w_6	w_7
User 1	1	0	0	0	1	1	0
User 2	0	1	1	0	0	1	1
User 3	1	1	0	0	1	0	0
User 4	0	1	0	1	0	0	1
User 5	1	1	0	0	1	1	0
User 6	0	1	0	1	0	1	1
User 7	1	0	1	0	1	0	0
User 8	1	0	0	1	1	1	0

(c)

Node	w_1	w_2	w_3	w_4	w_5	w_6	w_7
User 1	1	0	0	0	1	1	0
User 2	0	1	1	0	0	1	1
User 3	1	0	0	0	1	0	0
User 4	0	1	0	1	0	0	0
User 5	1	1	0	0	1	1	1
User 6	0	1	1	1	0	1	1
User 7	1	1	1	0	1	0	0
User 8	1	0	0	1	0	1	0

(d)

Node	w_1	w_2	w_3	w_4	w_5	w_6	w_7
User 1	0	1	1	0	0	1	1
User 2	0	1	1	0	0	1	1
User 3	1	1	1	0	1	1	1
User 4	0	0	0	1	0	0	1
User 5	0	1	1	0	0	1	1
User 6	0	1	1	0	0	1	1
User 7	0	1	1	0	0	1	1
User 8	0	1	1	0	0	1	1

(e)

Figure 3.8: (a) A sample network consisting of eight users, (b) each user has a profile with seven attributes with a weight vector, (c) initial sign flags for all eight users, (d) every user updates their strategy, and (e) the illustrated system converges.

CHAPTER 4

THEORETICAL ANALYSIS

In this section, we provide proofs for the theorems introduced and utilized in this dissertation.

Theorem 4.1. *Given that User x discloses k_x attributes and User y discloses k_y attributes, the probability of having exactly r pairs is given by*

$$P_{k_x,k_y}(N_p = r) = \frac{\binom{m}{k_x} \cdot \binom{k_x}{r} \cdot \binom{m-k_x}{k_y-r}}{\sum\limits_{i=\max\{k_x+k_y-m,0\}}^{\min\{k_x,k_y\}} \binom{m}{k_x} \cdot \binom{k_x}{i} \cdot \binom{m-k_x}{k_y-i}}.$$

Proof. User x has $\binom{m}{k_x}$ ways of selecting k_x attributes for disclosure from all m attributes in his/her profile. Similarly, there are $\binom{k_x}{r}$ ways of obtaining r pairs from the k_x disclosed attributes. Figure 3.1 shows a profile configuration arrangement in which users x and y have r pairs. As shown in the figure, to have exactly r pairs, the $k_x - r$ unpaired but revealed attributes of User x should not correspond with the $k_y - r$ unpaired but revealed attributes from User y. In fact, the $k_y - r$ unpaired but revealed attributes from User y can only correspond to the $m - k_x$ unrevealed attributes from User x.

Therefore, as shown in Equation 4.1, the number Δ_r of ways to obtain r pairs is a product of selecting k_x from m attributes, selecting r out of the k_x attributes, and selecting $k_y - r$ attributes from $m - k_x$ withheld attributes,

$$\Delta_r = \binom{m}{k_x} \cdot \binom{k_x}{r} \cdot \binom{m - k_x}{k_y - r}. \tag{4.1}$$

The number of pairs between two users N_p varies between two extremes. The highest possible number of pairs is achieved when there is maximum overlap between the attributes disclosed by both users. In this case, the number of pairs is equal to the smaller number of revealed attributes and is given by $N_p = \min\{k_x, k_y\}$. On the other hand, the lowest possible number of pairs is achieved when there is minimum overlap between the revealed attributes of both users. This number is equal to 0 if both sets of revealed attributes are completely disjointed, but equal to $k_x + k_y - m$ if the sum of the revealed attributes is higher than the number of profile attributes for any user. The maximum number of pairs is therefore given by $N_p = \max\{k_x + k_y - m, 0\}$. For example, if Alice reveals five out of seven attributes, and Bob reveals four out of the seven attributes, then the maximum number of pairs is four and the minimum number of pairs is two. However, if Alice reveals two of the seven attributes and Bob reveals four of the seven attributes, then the maximum and minimum number of pairs is two and zero, respectively.

Therefore, the number of pairs N_p is an integer in the range between $\max\{k_x + k_y - m, 0\}$ and $\min\{k_x, k_y\}$. The probability of having exactly r pairs is therefore obtained by dividing Δ_r by $\sum_i \Delta_i$ where i takes on all possible values of N_p. □

Theorem 4.2. *Given r pairs, the probability of getting η equal value pairs is*

$$P(N_{ep} = \eta \mid N_p = r) = \frac{(L-1)^{r-\eta}}{L^r} \cdot \binom{r}{\eta}.$$

Proof. We assume that each attribute can take one of the L possible values. Additionally, there exist r pairs between the two users x and y. To obtain η equal value pairs, η of the r pairs belonging to User x should match with User y, while the $r - \eta$ remaining paired attributes must exhibit one of the remaining $L - 1$ values not already exhibited by User y. Without any restrictions, there are a total of L^{2r} possible assignments for the r pairs. Given that η of the pairs have identical values, the numbers of possible assignments is given by $L^r \cdot 1^\eta \cdot (L-1)^{r-\eta}$. The probability of having η equal value pairs is a product of $\binom{r}{\eta}$ ways of selecting η equal value pairs from r pairs, and the ratio $\frac{L^r \cdot 1^\eta \cdot (L-1)^{r-\eta}}{L^{2r}}$. □

Theorem 4.3. *Given that User x and User y disclose k_x and k_y attributes, respectively, the defined negative utility function of User x in the two-user game is obtained by:*

$$R_{k_x,k_y} = \begin{cases} \frac{1}{1+\frac{\binom{m-k_x}{k_y-k_x}}{\binom{m}{k_y}} \frac{1}{L^{k_x}}} & \text{if } k_x \leq k_y, \\ 1 & \text{if } k_x > k_y. \end{cases}$$

Proof. The negative utility is given by the risk of identity inference. A user's identity is less likely to be inferred if that user is hidden by another user. In our model, the risk of User x is equated to the inverse of the number of users that exhibit the same characteristics (attributes and their values) exhibited by User x. For example, if Alice exhibits completely unique characteristics in a network, then her risk is 1. If, on the other hand, another user in the network exhibits the same characteristics as Alice (same name, age, hometown, etc.), then Alice's risk is $\frac{1}{2}$.

Formally, given two users x and y disclose k_x and k_y attributes, respectively, User x cannot be hidden by User y if $k_x > k_y$, because User x is more distinguishable than y since User x reveals more attributes that uniquely identify them compared to User y. In this case, the risk for User x takes the maximum value of 1 because the only user with x's characteristics is User x himself.

When $k_x \leq k_y$, there is a probability that User x is hidden by User y. User x is hidden by User y when all the attributes disclosed by User x are identical to attributes disclosed by User y. Given m attributes, there are a total of $\binom{m}{k_x}\binom{m}{k_y}$ ways for users x and y to select k_x and k_y attributes for revelation, respectively. For x to be hidden by y, the k_x revealed attributes should be equal value pairs and any extra attributes $(k_y - k_x)$ revealed by User y should correspond to the $m - k_x$ attributes that User x withheld. This can happen in $\binom{m}{k_x}\binom{m-k_x}{k_y-k_x}$ ways. Therefore, the probability that User x reveals attributes only among the attributes revealed by y is given by $\frac{\binom{m-k_x}{k_y-k_x}}{\binom{m}{k_y}}$. Assuming that any attribute can take up one of L possible values, the probability of x being hidden by y becomes $\frac{\binom{m-k_x}{k_y-k_x}}{\binom{m}{k_y}}\frac{1}{L^{k_x}}$. The number of users with x's characteristics is therefore $1 + \frac{\binom{m-k_x}{k_y-k_x}}{\binom{m}{k_y}}\frac{1}{L^{k_x}}$ and therefore User x's risk is given by $1/(1+\frac{\binom{m-k_x}{k_y-k_x}}{\binom{m}{k_y}}\frac{1}{L^{k_x}})$. $\square$

We consider the *risk-free* scenario by nullifying the influence of the risk factor (negative utility) on the utility function in Equation 3.6 and fitness equation in Equation 3.11. This is done by setting $R_{k_x,k_y} = 1$ so that the risk is unaffected by the number of revealed attributes of any users. In this way, the strategies of the users only vary with the positive utility and are therefore risk-free (or negative utility free).

CHAPTER 5

RESULTS AND DISCUSSION

5.1 Simulations Settings

We conduct simulations of the risk-included and risk-free cases for the two-user game, basic evolutionary game, and the weighted evolutionary game. In all games, we consider user profiles made up of seven attributes ($m = 7$), which each user can reveal or withhold. This is sufficiently large to make observations that can be applied to a generic online social network. We set the number of users in the basic evolutionary game to 80 so that all eight categories of users initially have a round number of members. We set the number of users in the weighted evolutionary game model to 100 to emphasize the differences in graph structure between the considered network topologies. Other simulation settings specific to particular games are provided below and in Table 5.1.

5.1.1 Two-user Game

In the two-user game, the motivation is set to $\alpha \in \{0.2\}$. For the risk-free cases, the risk $R_{i,j}$ is made independent of i and j by setting it to 1.

We use the payoff matrix to derive the Nash equilibria for the two-user game. The payoff matrix shows the payoff values for each strategy combination for both players. Each payoff value is calculated from Equation 3.6 and the resultant payoff

Table 5.1: Parameter Values Used in the Models

Parameter	**Value**		
	Two-user game	**Basic evolutionary game**	**Weighted evolutionary game**
α	0.2	0.2. 0.9. 1.5	–
$\lvert N\rvert$	2	80	100
$\lvert A\rvert$	7	7	7
w_P	–	–	1
w_N	–	–	15
W	–	–	(0.02. 0.06. 0.10. 0.14. 0.18. 0.22. 0.28)

matrices for the risk-included and risk-free two-user games are provided in Table 5.2 and Table 5.3, respectively.

Table 5.2: Payoff Matrix of the Two-user Game for Risk-included Scenario

y x	$k_y = 0$	$k_y = 1$	$k_y = 2$	$k_y = 3$	$k_y = 4$	$k_y = 5$	$k_y = 6$	$k_y = 7$
$k_x = 0$	0.1104. 0.1104	0.2209. 0.1104	0.2209. 0.1104	0.2209. 0.1104	0.2209. 0.1104	0.2209. 0.1104	0.2209. 0.1104	0.2209. 0.1104
$k_x = 1$	0.1104. 0.2209	0.0327. 0.0327	0.0717. 0.0654	0.1122. 0.0982	0.1558. 0.1309	0.2025. 0.1636	0.2524. 0.1963	0.1581. 0.1186
$k_x = 2$	0.1104. 0.2209	0.0654. 0.0717	0.0170. 0.0170	0.0517. 0.0509	0.1050. 0.1018	0.1787. 0.1697	0.1896. 0.1757	0.1415, 0.1273
$k_x = 3$	0.1104. 0.2209	0.0982. 0.1122	0.0509, 0.0517	0.0141. 0.0141	0.0566, 0.0564	0.1105. 0.1093	0.1541. 0.1509	0.1418. 0.1367
$k_x = 4$	0.1104. 0.2209	0.1309, 0.1558	0.1018. 0.1050	0.0564. 0.0566	0.0151. 0.0151	0.0588. 0.0587	0.1222. 0.1215	0.1486. 0.1468
$k_x = 5$	0.1104. 0.2209	0.1636. 0.2025	0.1697. 0.1787	0.1093. 0.1105	0.0587. 0.0588	0.0210. 0.0210	0.0871. 0.0870	0.1583. 0.1577
$k_x = 6$	0.1104. 0.2209	0.1963. 0.2524	0.1757. 0.1896	0.1509. 0.1541	0.1215. 0.1222	0.0870, 0.0871	0.0467. 0.0467	0.1695. 0.1693
$k_x = 7$	0.1104. 0.2209	0.1186, 0.1581	0.1273, 0.1415	0.1367. 0.1418	0.1468. 0.1486	0.1577. 0.1583	0.1693. 0.1695	0.1818. 0.1818

As an example, consider User x discloses three attributes, while User y discloses six attributes ($k_x = 3, k_y = 6$). Given $m = 7$, $\alpha = 0.2$, $\epsilon = 5.0$, and $L = 3$, the payoff value of the risk-included two-user game is calculated.

Table 5.3: Payoff Matrix of the Two-user Game for Risk-free Scenario

x \ y	$k_y = 0$	$k_y = 1$	$k_y = 2$	$k_y = 3$	$k_y = 4$	$k_y = 5$	$k_y = 6$	$k_y = 7$
$k_x = 0$	0.1104. 0.1104	0.1104. 0.1104	0.1104, 0.1104	0.1104. 0.1104	0.1104. 0.1104	0.1104, 0.1104	0.1104. 0.1104	0.1104. 0.1104
$k_x = 1$	0.1104. 0.1104	0.0327. 0.0327	0.0654. 0.0654	0.0982. 0.0982	0.1309. 0.1309	0.1636, 0.1636	0.1963. 0.1963	0.1186. 0.1186
$k_x = 2$	0.1104. 0.1104	0.0654. 0.0654	0.0170. 0.0170	0.0509. 0.0509	0.1018. 0.1018	0.1697. 0.1697	0.1757. 0.1757	0.1273. 0.1273
$k_x = 3$	0.1104. 0.1104	0.0982. 0.0982	0.0509. 0.0509	0.0141. 0.0141	0.0564. 0.0564	0.1093. 0.1093	0.1509. 0.1509	0.1367. 0.1367
$k_x = 4$	0.1104. 0.1104	0.1309. 0.1309	0.1018. 0.1018	0.0564. 0.0564	0.0151. 0.0151	0.0587. 0.0587	0.1215. 0.1215	0.1468. 0.1468
$k_x = 5$	0.1104. 0.1104	0.1636. 0.1636	0.1697. 0.1697	0.1093, 0.1093	0.0587. 0.0587	0.0210. 0.0210	0.0870. 0.0870	0.1577. 0.1577
$k_x = 6$	0.1104. 0.1104	0.1963. 0.1963	0.1757. 0.1757	0.1509. 0.1509	0.1215. 0.1215	0.0870. 0.0870	0.0467. 0.0467	0.1693. 0.1693
$k_x = 7$	0.1104. 0.1104	0.1186. 0.1186	0.1273. 0.1273	0.1367. 0.1367	0.1468. 0.1468	0.1577. 0.1577	0.1693. 0.1693	0.1818. 0.1818

User x's utility is calculated from Equation 3.6, $u_x = \frac{P_{3,6}(F)}{R_{3,6}}$, where $P_{3,6}(F)$ is calculated from Equation 3.4:

$$P_{3,6}(F) = \sum_{r=\max\{3+6-7,0\}}^{\min\{3,6\}} P_{3,6}(N_p = r) \cdot \sum_{\eta=0}^{r} P(F \mid N_{ep} = \eta) \cdot P(N_{ep} = \eta \mid N_p = r) = 0.1509,$$

while $R_{3,6}$ is calculated from Equation 3.5, $R_{3,6} = 1/(1+\frac{\binom{7-3}{6-3}}{\binom{7}{6}}\frac{1}{3^3}) = 0.9793$. User x's utility is therefore given by $u_x = \frac{0.1509}{0.9793} = 0.1541$. Similarly, User y's utility is evaluated from $u_y = \frac{P_{6,3}(F)}{R_{6,3}}$, where the positive utility $P_{6,3}(F) = P_{3,6}(F) = 0.1509$.

From Equation 3.5, we find that User y's negative utility is 1. This risk is maximum because User y cannot be hidden by User x. Therefore, User y's utility in the same scenario is $u_y = 0.1509$.

5.1.2 Basic Evolutionary Game

For the basic evolutionary game, we set the motivation $\alpha \in \{0.2, 0.9, 1.5\}$. Similar to the two-user game in Section 5.1.1, the negative utility $R_{i,j}$ in the risk-free case is made independent of i and j by setting it to 1. The social network considered consists of 80 users with eight possible strategies. The users are categorized according to the strategy they employ. After initializing the population in each of the eight categories to ten users, we observe how the populations of the categories change over time.

5.1.3 Weighted Evolutionary Game

In this section, we describe the underlying simulation settings. The simulations deal with risk-included and risk-free cases of the weighted evolutionary game.

The simulation is designed to consider user profiles with seven attributes ($m = 7$). Each user can choose to reveal or to withhold each of these attributes. A 7-bit flag is assigned to each user, which corresponds to the attributes. For example, the flag 1000110 for User 1 means that Attributes 1, 5 and 6 are revealed while Attributes 2, 3, 4, and 7 are withheld.

We begin by randomly assigning the attribute flag to all users of the network. During each iteration, each user has two options: maintain his/her attribute flag, or change it (by revealing or withholding a single attribute).

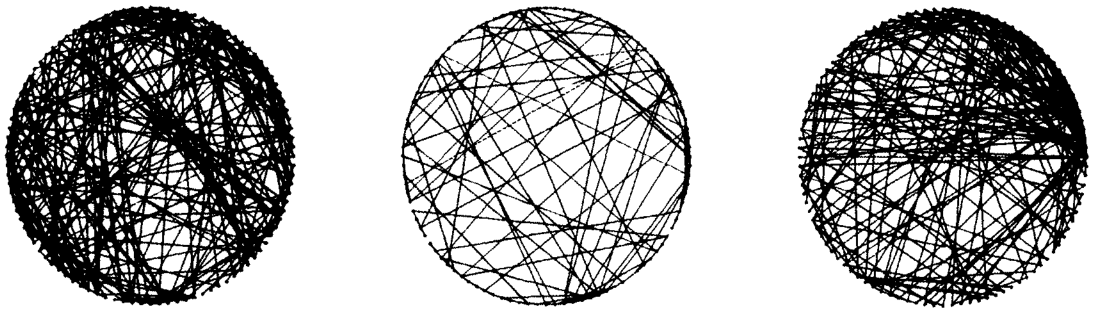
To consider different levels of the risk, we choose three different benefit-to-risk ratios (BRRs), which are $1 : 0$, $1 : 15$, and $1 : 30$ (cf. Table 5.1). While all the attributes are assigned to different weights, the weight vector for the attributes is assumed to be the same for each user of the network. Additional simulation settings are shown in Table 5.1. We run the simulation for each configuration 500 times. After averaging 500 simulation results, we obtain the dynamic curves in each of the considered networks, which include random, small-world, scale-free, and Facebook friend networks.

The size and average node degree for each network are all listed in Table 5.4. Random network, small-world network, and scale-free network are generated with the same size and average node degree. Two Facebook friend networks are collected from real Facebook accounts, which exhibit different size and average node degree.

Table 5.4: The Properties of Networks in the Simulation

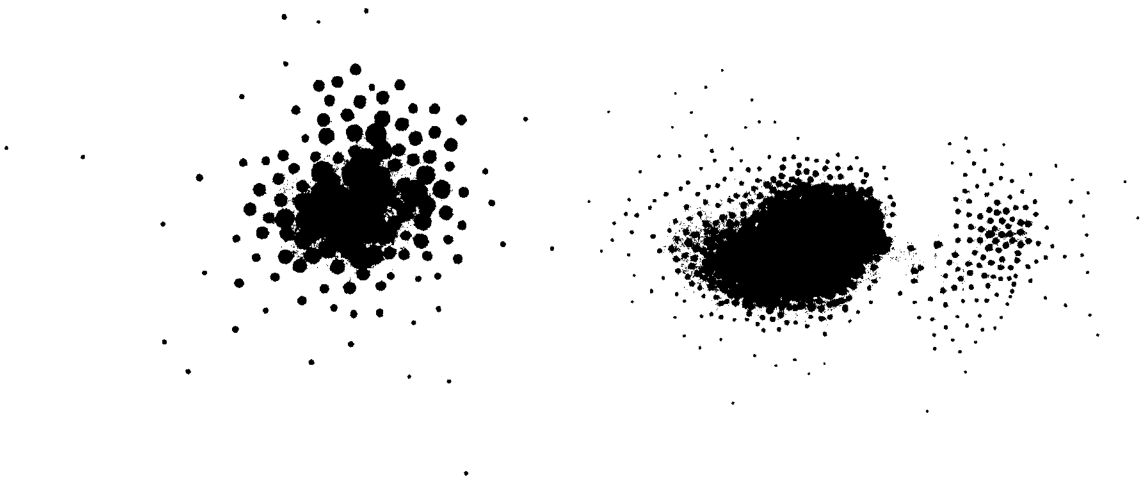
Network	**Size**	**Average Node Degree**
Random network	100	4
Small-world network	100	4
Scale-free network	100	4
FB1	151	15.0
FB2	502	49.0

In Figure 5.1, the visualized graphs for the random, small-world, and scale-free networks are shown. The visualized graphs for the Facebook friend networks are depicted in Figure 5.2. The Facebook graphs (FB1 and FB2) are obtained using the *SocialMediaData* function in *Mathematica.* Figure 5.2a and Figure 5.2b are from two different Facebook accounts.



(a) Random network **(b)** Small-world network **(c)** Scale-free network

Figure 5.1: The network topologies used in the simulations. The average node degree for each network is 4, and each network includes 100 nodes.



(a) FB1 (b) FB2

Figure 5.2: The Facebook friend networks used in the simulations. Network FB1 and FB2 are comprised of 151 and 502 nodes, respectively.

5.2 Results

5.2.1 Two-user Game

The Nash equilibria for the two user games in both the risk-included and risk-free scenarios are shown in Figure 5.3. The Nash equilibria correspond to the privacy decisions that the two users are likely to make. By definition, no user can increase his/her gain by changing his/her strategy unilaterally when Nash equilibrim is attained. The Nash equilibrium states are calculated using the *enumerating extreme* *points* method [44] as calculated by the Gambit tool [45]. The states represent all the possible final strategies employed by a risk-included or risk-free simulation. Figure 5.3 shows a total of 31 Nash equilibrium states, six of which are attained in the risk-free scenario, and 25 in the risk-included scenario. The states are represented with color-coded rectangles which correspond to the probability of a player taking that strategy. Darker colors correspond to higher probabilities while lighter colors correspond to lower probabilities. For example, State 25 of the risk-included scenario shows that both Player 1 and Player 2 choose strategy 1 with a probability 1.0. State 2 of the risk-included scenario shows that Player 1 could choose either strategy 7 or strategy 8 while Player 2 chooses strategy 1. Recall that Strategy i refers to the user's choice to reveal $i - 1$ attributes and therefore Player 1 choosing strategy 1 means they choose to withhold all their attributes.

The risk-included states show that Nash Equilibrium is only achieved when at least one of the players selects Strategy 1. All 25 equilibrium states involve at least one of the players employing Strategy 1, which corresponds to withholding all his/her attributes. This means that no player will be satisfied with his/her choice until at least one of them has chosen to withhold all their attributes. However, one player selecting to withhold all their attributes does not mean that the other player will choose the same strategy. For example, in State 2, Player 1 chooses to reveal six or seven of his seven attributes (strategy 7 or 8) with a probability of $\frac{628}{943}$ and $\frac{315}{943}$,s respectively.

The risk-free states show that players are more likely to reveal more attributes. Only one of the six equilibrium states has any player withholding any information

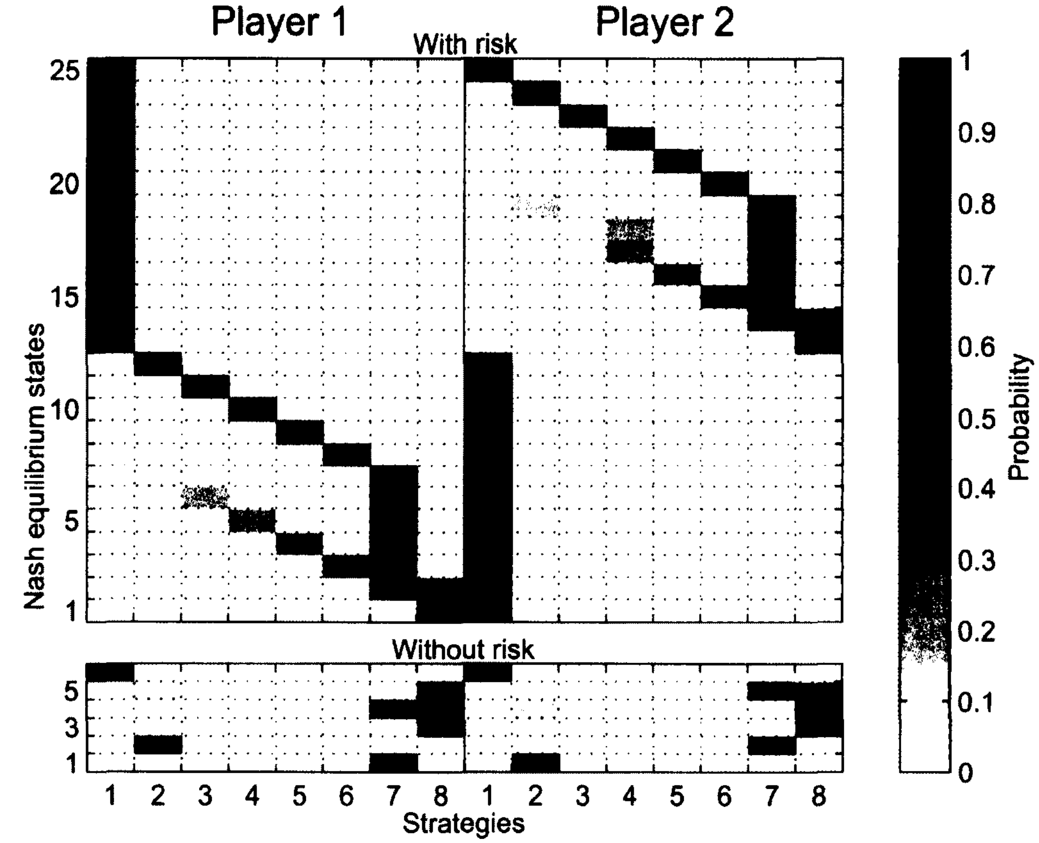


Figure 5.3: The correlation map shows the Nash equilibrium states for the two-player game with the colors corresponding to the probability of any user taking a specific strategy when Nash equilibrium is attained. Dark colors represent higher probabilities, and lighter colors represent lower probabilities.

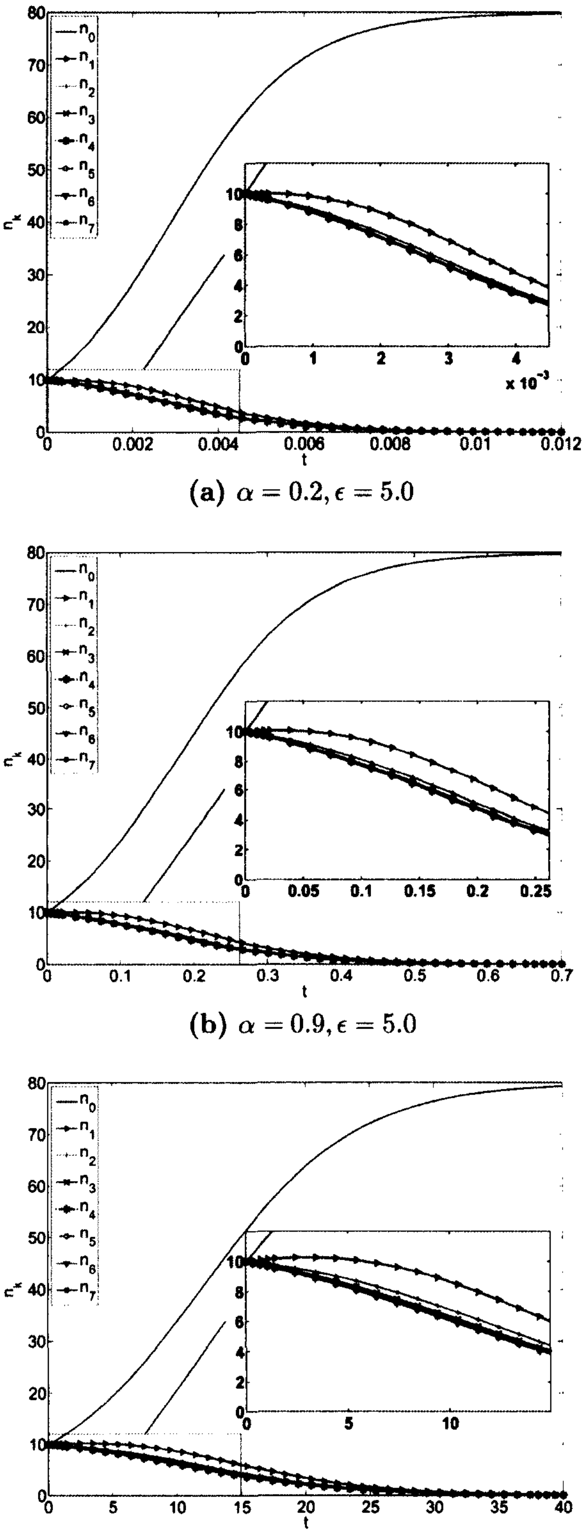
(State 6). All remaining five states have at least one of the players revealing a minimum of six out of seven attributes. For example, State 3 shows that both Player 1 and Player 2 choose Strategy 8, which is the reveal-all strategy.

The results show that with risk involved, equilibrium will only be attained if at least one of the players selects to withhold all his attributes. In a risk-free scenario, however, a player will only withhold all attributes if the other player has withheld all their attributes. Otherwise, both players are comfortable with revealing either all or all-but-one of their attributes.

5.2.2 Basic Evolutionary Game

The population dynamics for the basic evolutionary game model are shown in Figure 5.4 and Figure 5.5 for the risk-included and risk-free cases, respectively. The figures show how many users employ the different possible strategies and how this changes with time. By considering an 80-user network, and solving the system of differential equations for different values of motivation α, we are able to determine the number of users n_i that choose to reveal i attributes. For example, n_0 is the number of users who choose to withhold all their attributes (reveal 0 attributes), while n_7 is the number of users who choose to reveal all seven attributes. Initially, ten users are assigned to each strategy, i.e. $n_i = 10, \forall i \in \{0, 1, 2, 3, 4, 5, 6, 7\}$.

From Figure 5.4 we observe that the number of people who withhold all their attributes n_0 increases to 80 in all three plots, while all other strategies decrease to 0. This shows that as long as there is risk in the network, the final strategy employed by all users is to withhold all the attributes.

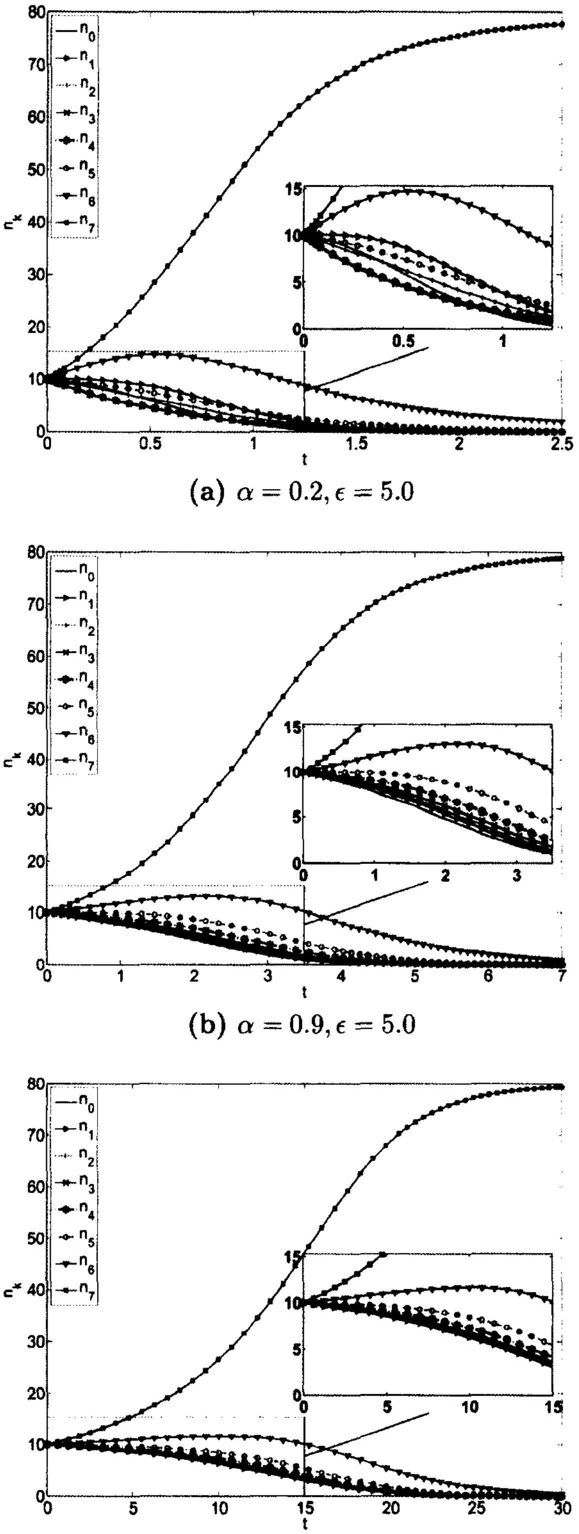


(a) $\alpha = 0.2, \epsilon = 5.0$

(b) $\alpha = 0.9, \epsilon = 5.0$

(c) $\alpha = 1.5, \epsilon = 5.0$

Figure 5.4: Population dynamics for the basic evolutionary game for the risk-included scenario with different levels of motivation.



(a) $\alpha = 0.2, \epsilon = 5.0$

(b) $\alpha = 0.9, \epsilon = 5.0$

(c) $\alpha = 1.5, \epsilon = 5.0$

Figure 5.5: Population dynamics for the basic evolutionary game for the risk-free scenario with different levels of motivation.

However, comparing Figures. 5.4a, 5.4b, and 5.4c shows that increasing the motivation α increases the time it takes for the network to eventually employ the withhold-all strategy. For example, comparing Figures. 5.4a and 5.4b shows that increasing α from 0.2 to 0.9 results in an increase in convergence time from 0.012 to 0.7. This indicates that an increase in the motivation to reveal more attributes only affects how long it takes for the network to eventually employ the withhold-all strategy. It is interesting to note that the reveal-one-attribute strategy n_1 is initially more common than the more revealing strategies $(n_2, n_3, n_4, n_5, n_6, n_7)$.

In contrast, we observe from Figure 5.5 that the dominant strategy in the risk-free case is to reveal all attributes n_7. The number of people who reveal all-but-one attribute n_6 also initially increases but eventually decreases alongside other less revealing strategies $(n_0, n_1, n_2, n_3, n_4, n_5)$. While this result might seem intuitive, the effect of increasing motivation is counter-intuitive. Figures. 5.5a, 5.5b, and 5.5c show that increasing the motivation increases the time the network takes to achieve equilibrium. For example, comparing Figures. 5.5a and 5.5b shows that increasing α from 0.2 to 0.9 results in an increase in convergence time from 2.5 to 7. Similar to the risk-included case, increasing motivation only affects the convergence time of the network. However, increasing the motivation also reduces the number of users who will initially employ the reveal all-but-one strategy n_6. This means that the *reveal all-but-one* strategy is more popular with lower values of motivation.

Comparing convergence times of the risk-free and risk-included networks shows that risk-free networks have longer convergence time than risk-included networks for low values of α, and a shorter convergence time for $\alpha = 1.5$.

5.2.3 Weighted Evolutionary Model

In this section, we describe the results derived from simulations of the weighted evolutionary game on a random network, a small-world network, a scale-free network, and two Facebook friend networks.

The attribute dynamic curves for random network, small-world network, scale-free network, FB1, and FB2 are shown in Figures. 5.6 - 5.15, respectively. Each dynamic curve shows how the proportion of the entire population that discloses any specific attribute changes with time. Each dynamic plot consists of seven curves corresponding to seven attributes, *Attr*$\#1$, *Attr*$\#2$, ..., *Attr*$\#7$, which are numbered according to their importance (weight), i.e. *Attr*$\#7$ is the most important attribute, while *Attr*$\#1$ is the least important attribute.

There are three sub-figures (Figures. a-c) in each figure. Figures. 5.6, 5.8, 5.10, 5.12, and 5.14 correspond to the simulation results when we consider the benefit and risk only within the users' friends. Figures. 5.7, 5.9, 5.11, 5.13, and 5.15 correspond to simulation results when we consider both the users' friends and friends-of-friends. The top, middle and bottom rows correspond to BRR values of $1:0$, $1:15$, and $1:30$, respectively, where $(BRR = 1:0)$ represents risk-free scenario.

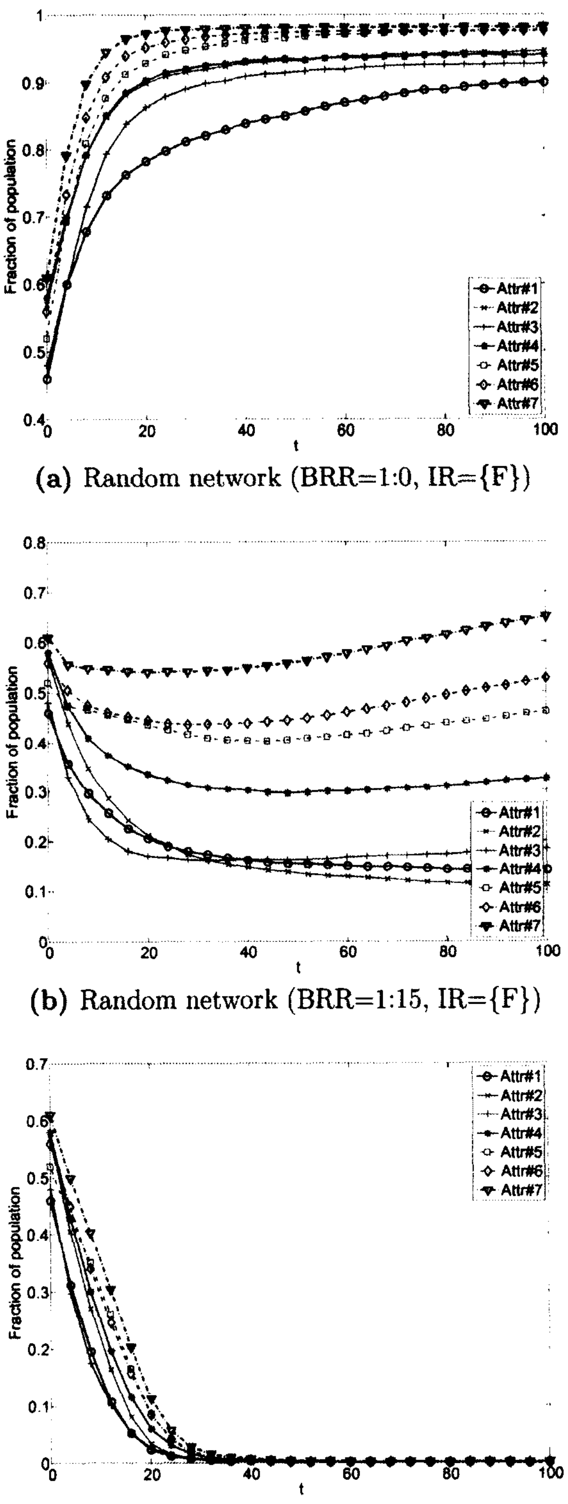
The first observation is a general reduction in attribute revelation with an increase in risk. Consider Figure 5.6 which shows the attribute dynamics in a random network: comparing Figures. 5.6a, 5.6b, and 5.6c shows that increasing the risk causes less users to reveal attributes. Figure 5.6a shows that over 85% of the population reveal all their attributes by 100 iterations when there is no risk. Introducing risk causes users to reveal less attributes. In fact, Figure 5.6c shows that all users withhold all their attributes by 50 iterations when $BRR = 1:30$. The small-world, scale-free, and Facebook networks (cf. Figures. 5.8, 5.10, 5.12, and 5.14) all exhibit similar observations. While this observation might seem intuitive, it provides some form of vindication for our model.

The second observation is that the networks generally exhibit larger drops in attribute revelation when the range of influence is restricted to friends as opposed to when friends-of-friends are also considered. For example, Figures. 5.8a and 5.9a show almost identical levels of revelation without risk. However, increasing the risk leads to more attributes withholding in Figures. 5.8b and 5.8c than it does in Figures. 5.9b and 5.9c. This means that risk plays a more dominant role in attribute disclosure when only the friends of a user are considered.

The third observation is that increasing the users' range of influence generally results in increased levels of attribute revelation. Consider Figures. 5.10 and 5.11 which capture attribute dynamics in a scale-free network: comparing Figures. 5.10a, 5.10b, and 5.10c and Figures. 5.11a, 5.11b, and 5.11c shows that maximum revelation is obtained by as early as 40 iterations for all attributes when friends-of-friends are considered (Figures. 5.11a, 5.11b, 5.11c). In contrast, the risk-free scenario with friends (Figure 5.10a) only obtains maximum revelation for some of the attributes, while Figures. 5.10b and 5.10c do not obtain maximum revelation for any attributes at all. This observation can be attributed to the process of enlarging the influential range. Increasing the range results in an increase in the number of users who can hide any specific user which leads to a reduction in risk. Increasing the range also allows for more users who share the same attributes, which leads to an increase in the user's benefit.

The next observation is related to the friend's influential range. Increasing the risk factor has a larger effect on attribute disclosure in the random and small-world networks than in the scale-free and Facebook networks. Comparing Figures. 5.6 and 5.8 to Figures. 5.10, 5.12 and 5.14 shows that $BRR = 1 : 30$ causes complete attribute withholding in the random and small-world networks in contrast to partial attribute withholding in the scale-free and Facebook networks.

The final observation is related to the effect of network topology on attribute disclosure with the range of influence restricted to friends. We find that network topology plays a more considerable effect on the privacy in risk-included scenarios than in a risk-free scenario for the random, small-world, and scale-free networks. Comparing Figures. 5.6a, 5.8a and 5.10a shows that the networks exhibit similar performance in the risk-free environment ($BRR = 1 : 0$). However, comparing Figures. 5.6b, 5.8b and 5.10b as well as Figures. 5.6c, 5.8c and 5.10c shows that the performance is different for different networks. For example, Figures. 5.6c and 5.8c show complete attribute withholding while Figure 5.10c shows partial attribute disclosure.

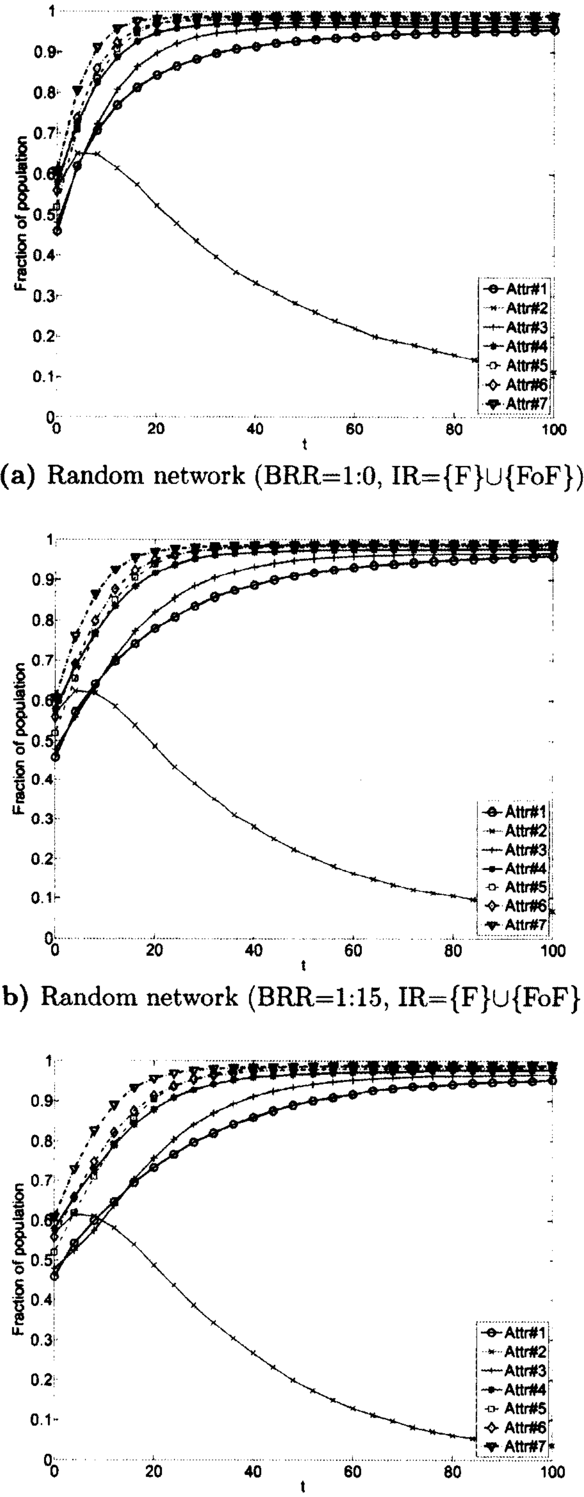


(a) Random network (BRR=1:0, IR={F})

(b) Random network (BRR=1:15, IR={F})

(c) Random network (BRR=1:30, IR={F})

Figure 5.6: Attribute dynamics for the weighted evolutionary game in random network, where the influential range of the utility function includes Friends.

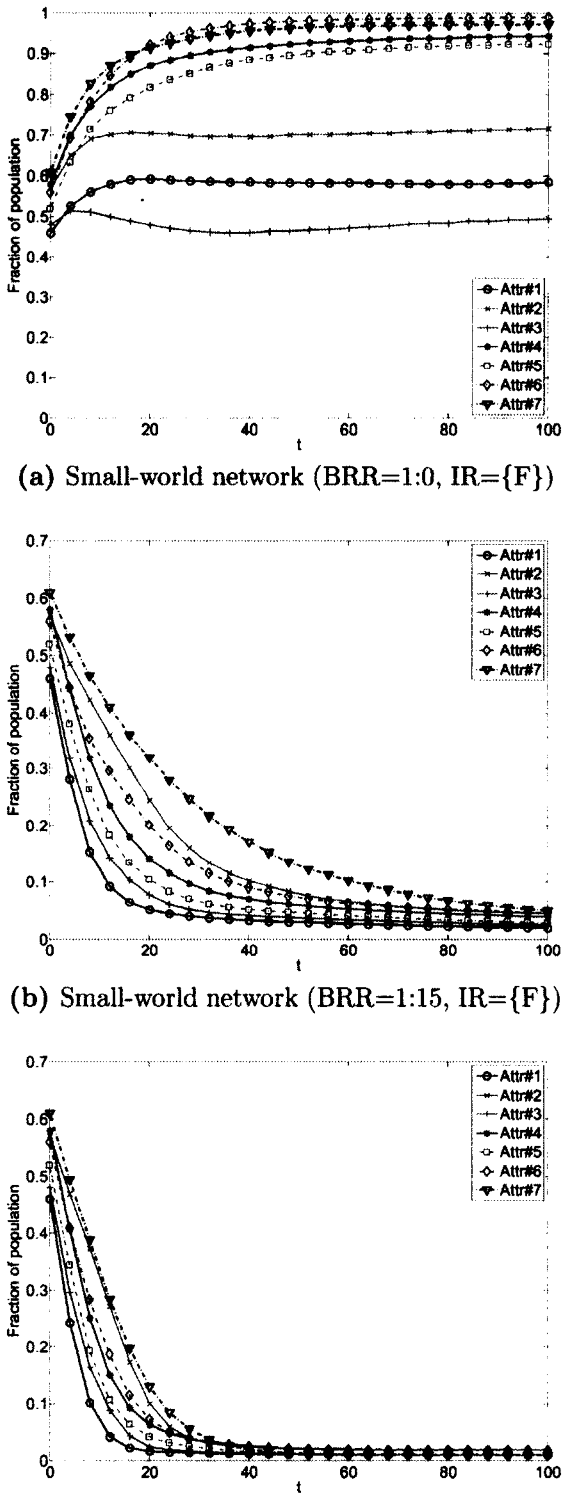


(a) Random network (BRR=1:0, IR={F}∪{FoF})

(b) Random network (BRR=1:15, IR={F}∪{FoF})

(c) Random network (BRR=1:30, IR={F}∪{FoF})

Figure 5.7: Attribute dynamics for the weighted evolutionary game in random network, where the influential range of the utility function includes Friends and Friends-of-Friends.

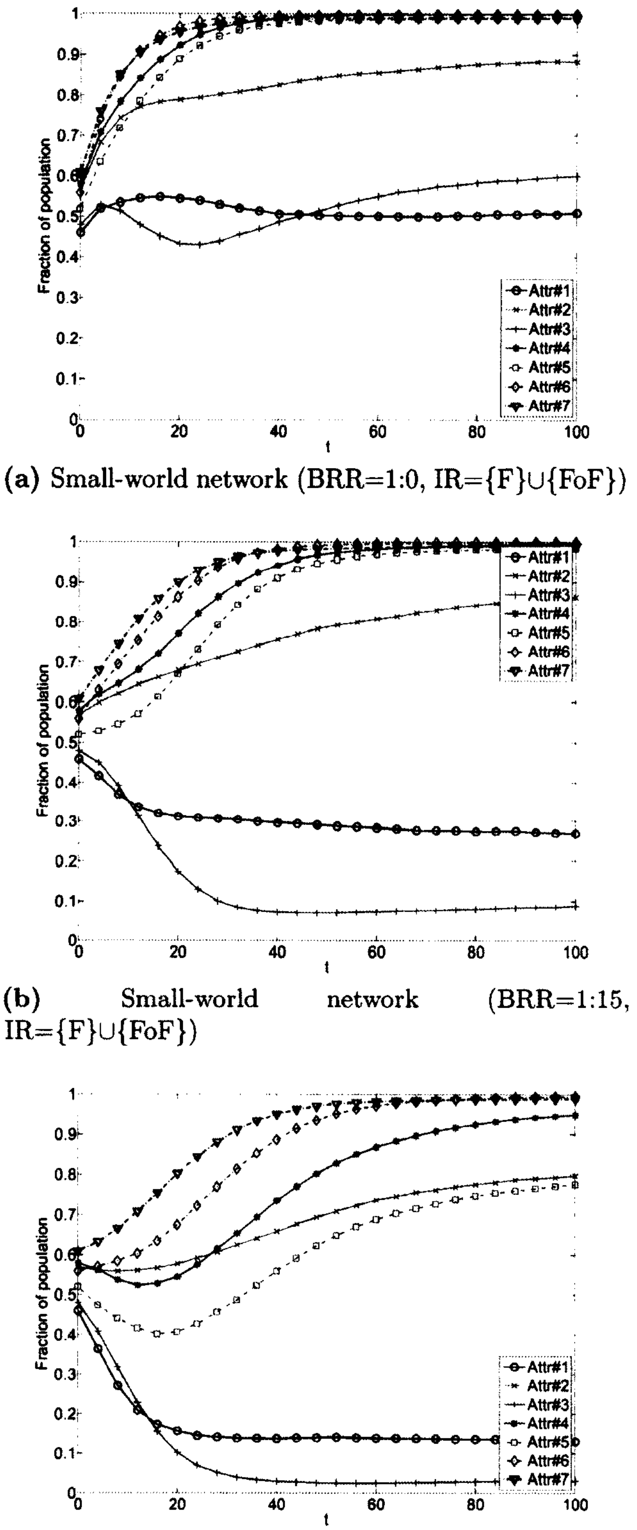


(a) Small-world network (BRR=1:0, IR={F})

(b) Small-world network (BRR=1:15, IR={F})

(c) Small-world network (BRR=1:30, IR={F})

Figure 5.8: Attribute dynamics for the weighted evolutionary game in small-world network, where the influential range of the utility function includes Friends.

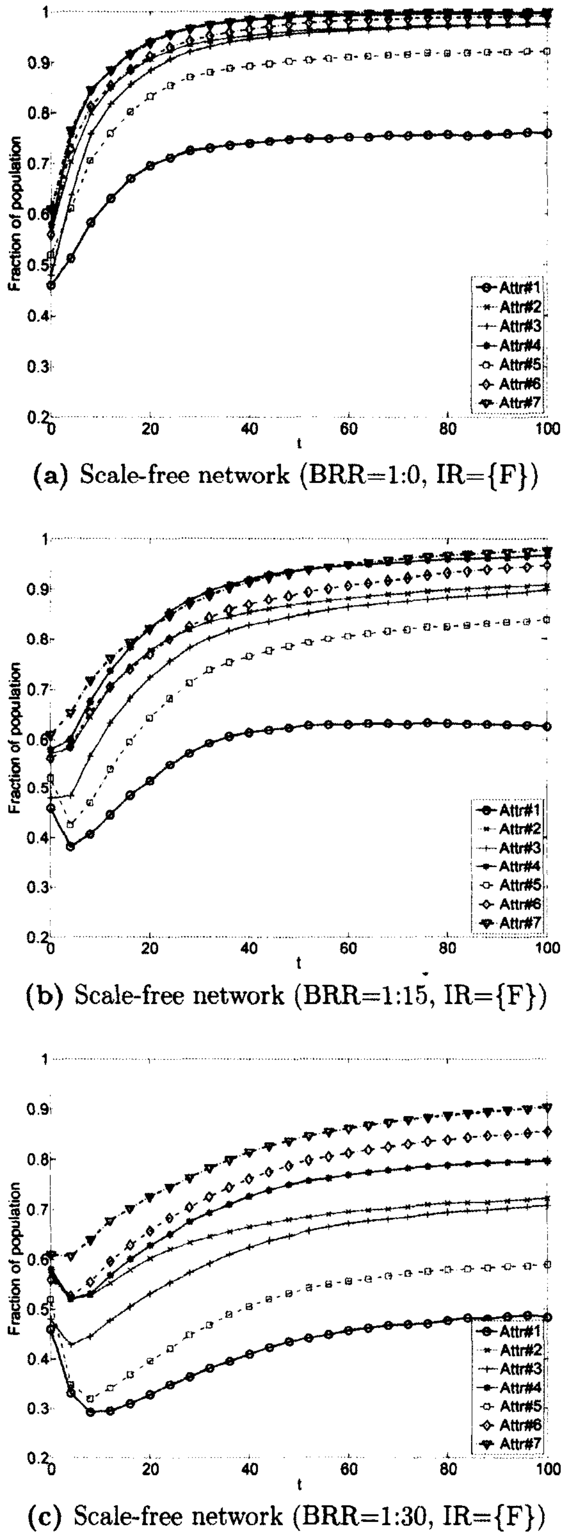


(a) Small-world network (BRR=1:0, IR={F}∪{FoF})

(b) Small-world network (BRR=1:15, IR={F}∪{FoF})

(c) Small-world network (BRR=1:30, IR={F}∪{FoF})

Figure 5.9: Attribute dynamics for the weighted evolutionary game in small-world network, where the influential range of the utility function includes Friends and Friends-of-Friends.

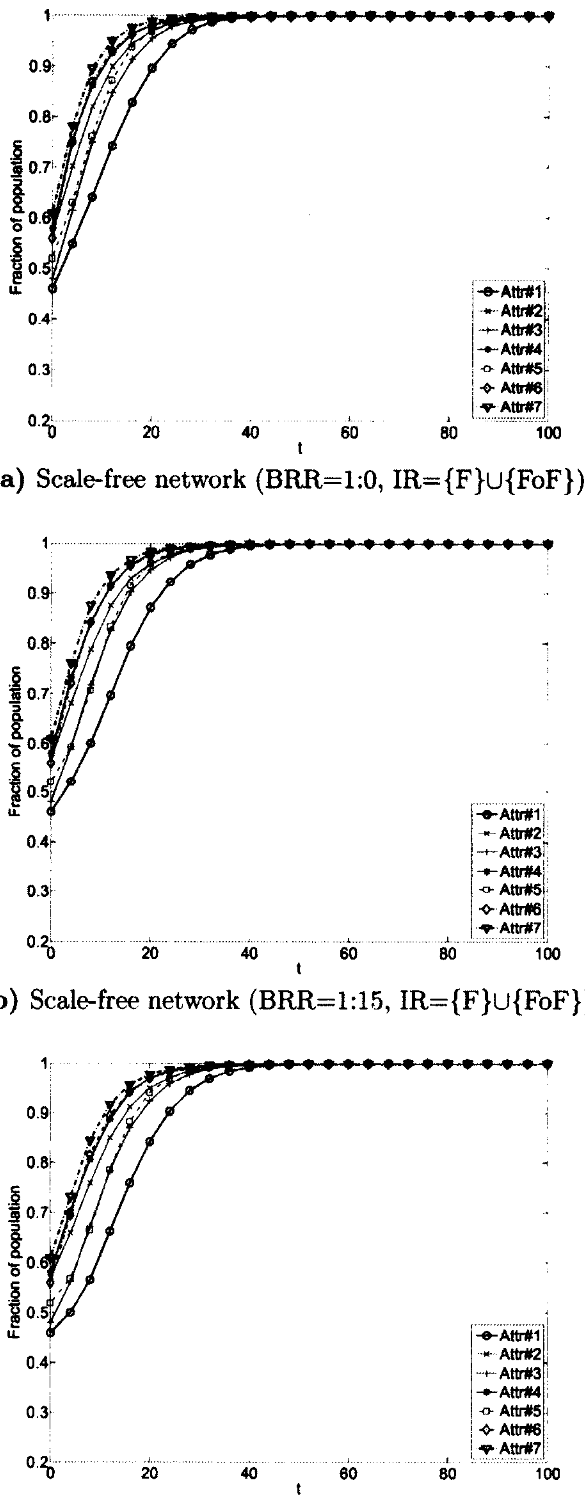


(a) Scale-free network (BRR=1:0, IR={F})

(b) Scale-free network (BRR=1:15, IR={F})

(c) Scale-free network (BRR=1:30, IR={F})

Figure 5.10: Attribute dynamics for the weighted evolutionary game in scale-free network, where the influential range of the utility function includes Friends.

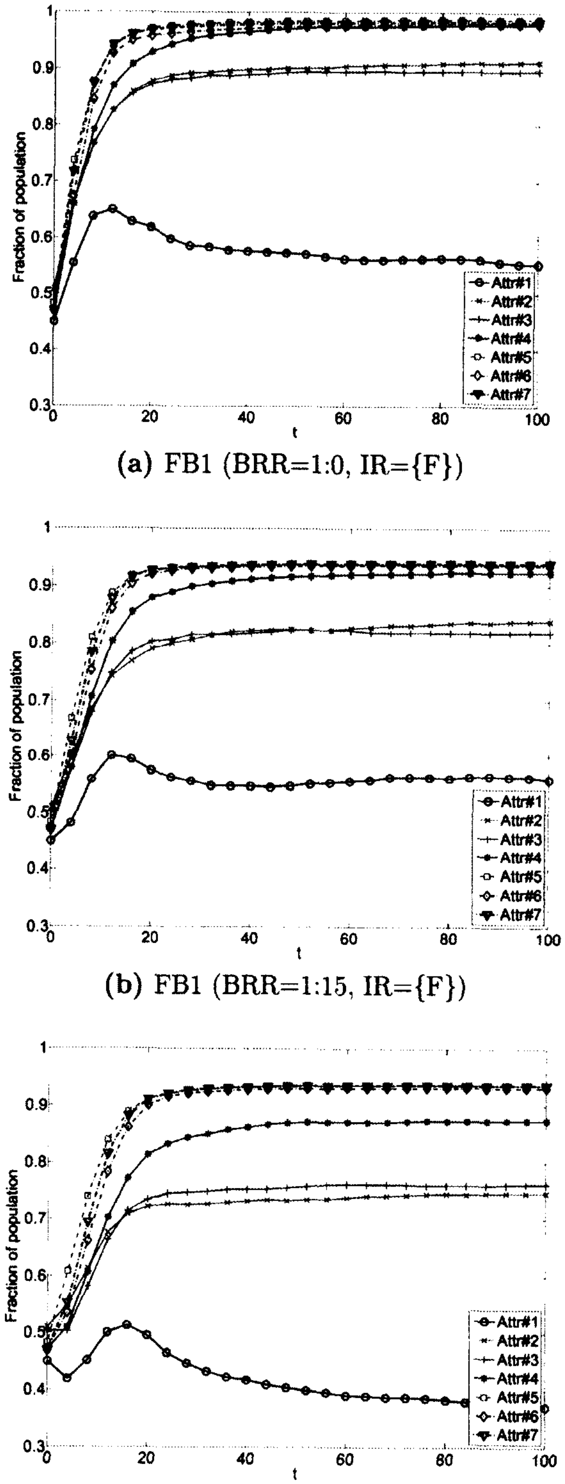


(a) Scale-free network (BRR=1:0, IR={F}∪{FoF})

(b) Scale-free network (BRR=1:15, IR={F}∪{FoF})

(c) Scale-free network (BRR=1:30, IR={F}∪{FoF})

Figure 5.11: Attribute dynamics for the weighted evolutionary game in scale-free network, where the influential range of the utility function includes Friends and Friends-of-Friends.

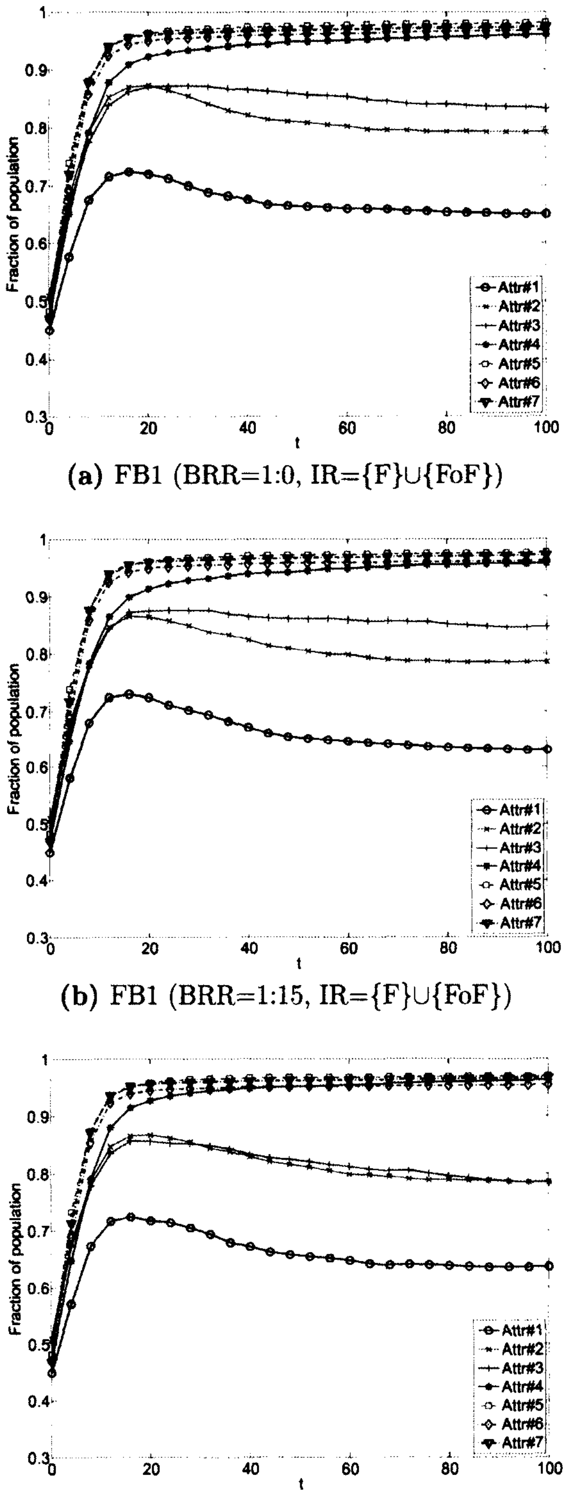


(a) FB1 (BRR=1:0, IR={F})

(b) FB1 (BRR=1:15, IR={F})

(c) FB1 (BRR=1:30, IR={F})

Figure 5.12: Attribute dynamics for the weighted evolutionary game in FB1, where the influential range of the utility function includes Friends.

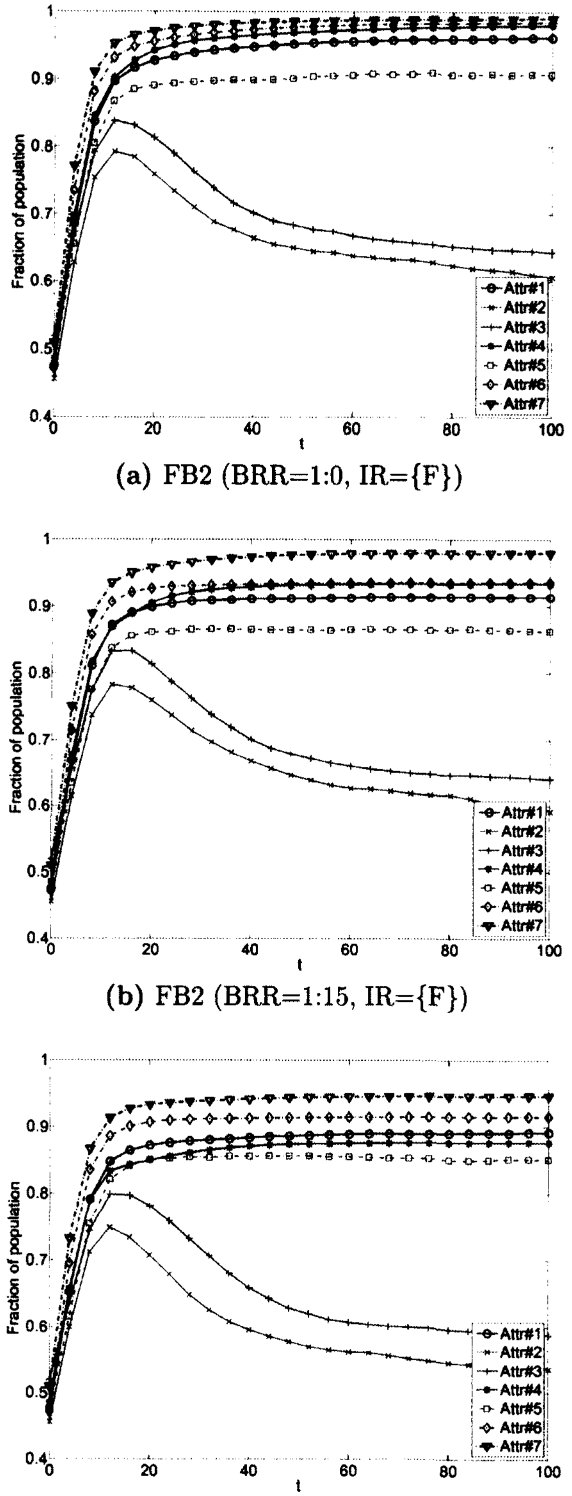


(a) FB1 (BRR=1:0, IR={F}∪{FoF})

(b) FB1 (BRR=1:15, IR={F}∪{FoF})

(c) FB1 (BRR=1:30, IR={F}∪{FoF})

Figure 5.13: Attribute dynamics for the weighted evolutionary game in FB1, where the influential range of the utility function includes Friends and Friends-of-Friends.

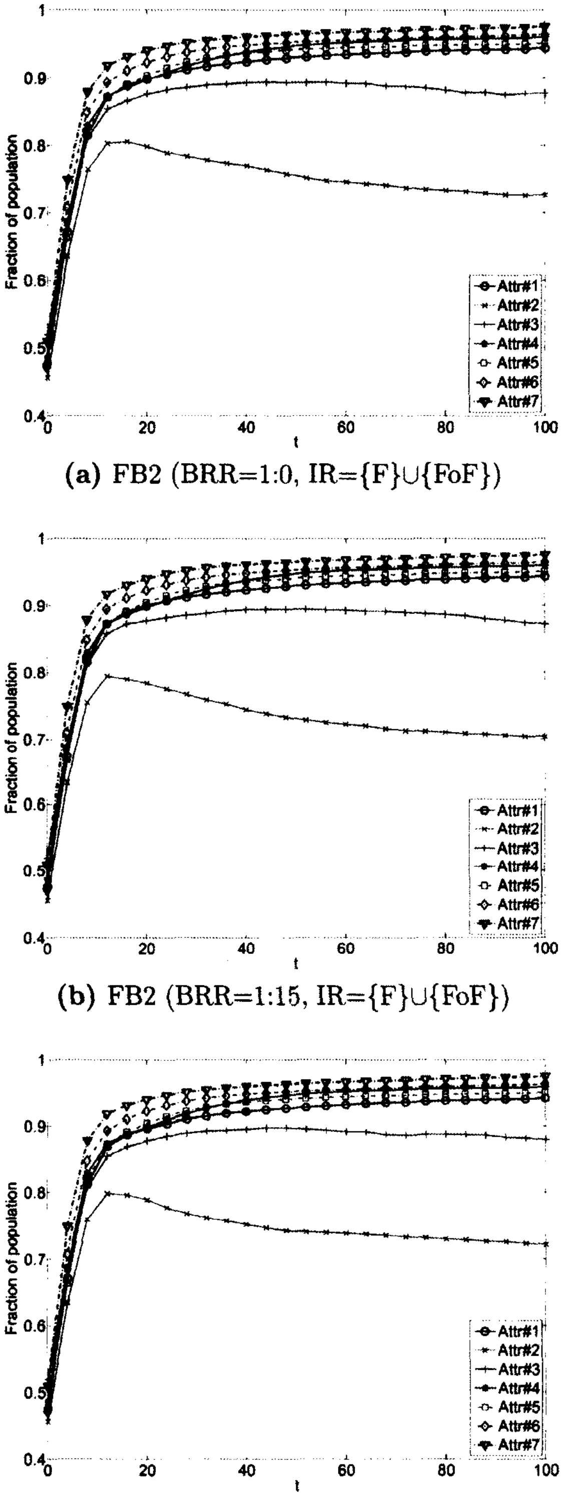


(a) FB2 (BRR=1:0, IR={F})

(b) FB2 (BRR=1:15, IR={F})

(c) FB2 (BRR=1:30, IR={F})

Figure 5.14: Attribute dynamics for the weighted evolutionary game in FB2, where the influential range of the utility function includes Friends.



(a) FB2 (BRR=1:0, IR={F}∪{FoF})

(b) FB2 (BRR=1:15, IR={F}∪{FoF})

(c) FB2 (BRR=1:30, IR={F}∪{FoF})

Figure 5.15: Attribute dynamics for the weighted evolutionary game in FB2, where the influential range of the utility function includes Friends and Friends-of-Friends.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

In this dissertation, we model and analyze the privacy settings of social networks from a game-theoretic perspective by introducing three types of game models. The models aim to investigate the influence of various factors on the privacy settings employed in social networks. The first is a two-user game model, which investigates the relationship between two users when disclosing profile attributes. The second is the basic evolutionary game model, which shows the dynamic behavior of multiple users as in large-scale online social networks. The third is a weighted evolutionary game model. We aim to investigate the influence of various factors such as attribute importance, benefit, risk, and network topology on the privacy settings employed in social networks.

As for results, we find both pure and mixed Nash equilibria in the two-user game. We also show the dominant strategy and population dynamics for the basic and weighted evolutionary game models in both the risk-included and risk-free cases. The two-user model results show that in a risk-included environment, Nash equilibrium is only achieved when at least one of the users withholds all their attributes. In the risk-free environment, results show that the ultimate privacy settings selected by one user is highly dependent on the privacy settings selected by another user.

In the basic evolutionary model, results indicate that the presence or absence of risk affects the final strategy adopted by the users in a network, while motivation only affects how long they take to adopt that strategy. This means that increasing motivation in a social network, e.g. by improving friend recommendation algorithms, only affects the level of self-disclosure in the short term. The existence of any risk factor means that eventually all users of a social network will adopt the highest possible privacy regardless of the benefits of revealing more profile attributes.

In the weighted evolutionary model, the results show that the most important attributes exhibit higher levels of revelation than the least important attributes. This finding is more evident in random and scale-free networks than in small-world networks. We also find that increasing the risk exhibits limited effect on the privacy dynamics of the network if we consider the benefit and risk from friends-of-friends. In the Facebook friend networks, which include more users and feature a higher average node degree, increasing the risk coefficient only slightly affects the level of attribute disclosure.

The models presented in this dissertation provide a way to study and comprehend the dynamics of privacy settings in social networks. Additionally, the nature of the transitions reveals the influence of certain factors in the short and long run in social network privacy to social network designers and users.

For future work, we plan to investigate the performance of our model on a larger variety of networks as well as compare it with data from real world social networks. Moreover, we intend to investigate multi-level privacy where users reveal different sets of attributes to different users in the network.

BIBLIOGRAPHY

[1] Guynn J: New Facebook information sharing features cause privacy concerns. http://articles.latimes.com/2011/sep/27/business/la-fi-facebook-privacy-20110927, September 27, 2011.

[2] Sengupta S: F.T.C. Settles Privacy Issue at Facebook. http://www.nytimes.com/2011/11/30/technology/facebook-agrees-to-ftc-settlement-on-privacy.html, November 29, 2011.

[3] Ben Abdesslem F, Parris I, Henderson T: Reliable Online Social Network Data Collection. In Computational Social Networks, London, UK: Springer, 2012:183-210.

[4] Scott C: Facebook Proposes More Changes to Privacy Policy. http://www.pcworld.com/businesscenter/article/255518/facebook_proposes_more_changes_to_privacy_policy.html, May 11, 2012.

[5] Mike Shor: Glossary of game theory terms. http://www.gametheory.net/dictionary/, May 6, 2002.

[6] Jundong Chen, Matthias R. Brust, Ankunda R. Kiremire, Vir V. Phoha: Modeling Privacy Settings of an Online Social Network from a Game-Theoretical Perspective. Proc. of 9th IEEE Int'l Conf. on Collaborative Computing, pages 213-220, Austin, TX, USA, 2013.

[7] Jundong Chen, Matthias R. Brust, Ankunda R. Kiremire, Vir V. Phoha: A Game Theoretic Approach for Modeling Privacy Settings of an Online Social Network. EAI Endorsed Transactions on Collaborative Computing, 2014(1):e4.

[8] Jundong Chen, Matthias R. Brust, Ankunda R. Kiremire, Vir V. Phoha: Modeling online social network users' profile attribute disclosure behavior from a game theoretic perspective, Comput. Commun., 2014(49):18-32.

[9] Sarah Evans: Top 18 social networks who have joined the 100 million (and more) users club. http://sarahsfav.es/2013/04/24/socialnetworks/, April 24, 2013.

[10] Pingdom: Facebook's crushing domination - the 26 busiest social networks. http://royal.pingdom.com/2012/07/25/facebooks-crushing-domination-the-26-busiest-social-networks/, July 25, 2012.

[11] University of Wyoming: Learning Guide: Pinterest. http://www.wyomingextension.org/wiki/index.php5?title=Learning_Guide:Pinterest, May 27, 2013.

[12] Goodreads: How it works. http://www.goodreads.com/about/how_it_works, May 27, 2013.

[13] Jon Kleinberg, Siddharth Suri, Eva Tardos, Tom Wexler: Strategic Network Formation with Structural Holes. Proc. of ACM SIGecom Exch., pages 1-4, Chicago, IL, USA, 2008.

[14] Osborne, Martin J.: An introduction to game theory. Oxford Univ. Press., 2004.

[15] Wei Chen, Zhenming Liu, Xiaorui Sun, Yajun Wang: Community Detection in Social Networks Through Community Formation Games. 22nd International Joint Conference on Artificial Intelligence (IJCAI), pages 2576-2581, Menlo Park, CA, USA, 2011.

[16] R. Narayanam, Y. Narahari: A Shapley Value-Based Approach to Discover Influential Nodes in Social Networks. IEEE Trans. Autom. Sci. Eng. 2011, 8(1):130-147.

[17] Andrew K. C. Wong, David E. Ghahraman: Random Graphs: Structural-Contextual Dichotomy. IEEE Trans. Pattern Anal. Mach. Intell., 1980, 2(4):341-348.

[18] P. Berenbrink, A. Brinkmann, T. Friedetzky, L. Nagel: Balls into non-uniform bins. IEEE International Symposium on Parallel Distributed Processing (IPDPS), pages 1-10, Atlanta, GA, USA, 2010.

[19] Liu Y, Gummadi KP, Krishnamurthy B, Mislove A: Analyzing Facebook privacy settings: User expectations vs. reality. Proc. of the Usenix/ACM Internet Measurement Conf. (IMC), pages 61-70, Berlin, Germany, 2011.

[20] Madden M: Privacy management on social media sites. `http://pewinternet.org/Reports/2012/Privacy-management-on-social-media/Summary-of-findings.aspx`, 2012.

[21] Krasnova H, Spiekermann S, Koroleva K, Hildebrand T: Online social networks: why we disclose. Journal of Information Technology, 2010, 25:109-125.

[22] Miller M, Lovin LS, Cook JM: Birds of a Feather: Homophily in Social Networks. Annual Review of Sociology, 2001, 27:415-444.

[23] Kossinets G, Watts DJ: Origins of Homophily in an Evolving Social Network. American Journal of Sociology, 2009, 115(2):405-450.

[24] Lampe C, Ellison N, Steinfield C: A familiar Face(book): Profile elements as signals in an online social network. Proc. of the SIGCHI Conf. on Human Factors in Computing Systems, pages 435-444, San Jose, CA, USA, 2007.

[25] Mislove A, Viswanath B, Gummadi KP, Druschel P: You Are Who You Know: Inferring User Profiles in Online Social Networks. Proc. of the ACM Int'l Conf. on Web Search and Data Mining, pages 251-260, Hong Kong, China, 2010.

[26] Becker J, Chen H: Measuring privacy risk in online social networks. Web 2.0 Security and Privacy Workshop, Oakland, CA, 2009.

[27] Squicciarini AC, Griffin C, Sundareswaran S: Towards a Game Theoretical Model for Identity Validation in Social Network Sites. Proc. of the Third Int'l Conf. on Social Computing, pages 1081-1088, Boston, MA, USA, 2011.

[28] Squicciarini AC, Griffin C: An Informed Model of Personal Information Release in Social Networking Sites. In ASE/IEEE Conf. on Privacy, Security, Risk and Trust, pages 635-645, Amsterdam, Netherlands, 2012.

[29] Skyrms B: The Stag Hunt and the Evolution of Social Structure. Cambridge University Press, 2003.

[30] Lin W, Zhao H, Liu K: Cooperation Stimulation Strategies for Peer-to-Peer Wireless Live Video- Sharing Social Networks. IEEE Trans. Image Processing, 2010, 19(7):1768-1784.

[31] Yu W, Liu K: Secure Cooperation in Autonomous Mobile Ad-Hoc Networks Under Noise and Imperfect Monitoring: A Game-Theoretic Approach. IEEE Trans. Inf. Forensics Security, 2008, 3(2):317-330.

[32] Xu S, Li X, Parker T, Wang X: Exploiting Trust-Based Social Networks for Distributed Protection of Sensitive Data. IEEE Trans. Inf. Forensics Security, 2011, 6(1):39-52.

[33] Squicciarini AC, Shehab M, Paci F: Collective Privacy Management in Social Networks. Proc. of the Int'l World Wide Web Conf., pages 521-530, Madrid, Spain, 2009.

[34] Kamhoua C, Pissinou N, Makki K: Game Theoretic Modeling and Evolution of Trust in Autonomous Multi-hop Networks: Application to Network Security and Privacy. Proc. of the IEEE Int'l Conference on Comms. (ICC), pages 1-6, Kyoto, Japan, 2011.

[35] Antonioni A, Tomassini M: Cooperation on Social Networks and its Robustness. Advances in Complex Systems, 2012(15):46-64.

[36] Osborne MJ, Rubinstein A: A course in game theory. Cambridge, MA: The MIT Press, 1994.

[37] Wong AKC, Ghahraman DE: Random Graphs: Structural-Contextual Dichotomy. Trans. Pattern Anal. Machine Intell., 1980, 2(4):341-348.

[38] Newman MEJ, Watts DJ, Strogatz SH: Random graph models of social networks. Proc. Natl. Acad. Sci. U.S.A., 2002, 99(1):2566-2572.

[39] Erdös P, Renyi A: On the evolution of random graphs. Publ. Math. Inst. Hung. Acad. Sci., 1960, 5:17-61.

[40] Watts DJ, Strogatz SH: Collective dynamics of small-world networks. Nature, 1998, 393(6684):440-442.

[41] Barabási AL, Albert R: Emergence of Scaling in Random Networks. Science, 1999, 286(5439):509-512.

[42] Taylor P, Jonker L: Evolutionary Stable Strategies and Game Dynamics. Mathematical Biosciences, 1978, 40(1-2):145-156.

[43] Roca CP, Cuesta JA, Sánchez A: Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics. Phys. Life Rev., 2009, 6(4):208-249.

[44] Audet C, S, Hansen P: Enumeration of All Extreme Equilibria in Game Theory: Bimatrix and Polymatrix Games. Journal of Optimization Theory and Applications, 2006, 129(3):349-372.

[45] McKelvey, D R, McLennan, M A, Turocy TL: Gambit: Software Tools for Game Theory. `http://www.gambit-project.org`, 2010.