GROWTH MODEL OF ONLINE FRIENDSHIP AND INDIVIDUAL, DYADIC, AND NETWORK DETERMINANTS

ZHANG LUN

DOCTOR OF PHILOSOPHY

CITY UNIVERSITY OF HONG KONG

SEPTEMBER 2011
Growth Model of Online Friendship and Individual, Dyadic, And Network Determinants

Submitted to
Department of Media and Communication
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

By

Zhang Lun

September 2011
Contents

Contents................................................................................................................................................. i

Abstract.................................................................................................................................................. iv

Acknowledgement................................................................................................................................... vii

List of Tables........................................................................................................................................ ix

List of Figures.......................................................................................................................................... x

Chapter 1. Introduction.......................................................................................................................... 1

Chapter 2. Friendship Formation and Network Evolution: Theories and Models ......................... 5

2.1. Theoretical Perspective of Friendship Formation ......................................................................... 5

2.1.1 Balance theory................................................................................................................................. 5

2.2.2 Homophily ..................................................................................................................................... 7

2.2.3 Rational choice theory, social exchange theory, and reciprocity principle ............................. 8

2.2.4 Preferential attachment.................................................................................................................. 10

2.2 The Directionality of Relationships Indicated by the Theories...................................................... 11

2.3 Empirical Studies of Friendship Formation: A Grand Picture .................................................. 14

2.3.1 Context........................................................................................................................................ 15

2.3.2 Theoretical perspectives................................................................................................................ 15

2.3.3 Self-reported measurement of personal social network.............................................................. 17

2.3.4 Sample size................................................................................................................................ 18

2.3.5 Methods...................................................................................................................................... 19

2.3.6 Summary..................................................................................................................................... 20

2.4 Models of Network Evolution........................................................................................................ 21
2.4.1 Simulation model ........................................................................................................ 22
2.4.2 Causality testing approach .......................................................................................... 25
2.4.3 Summary ..................................................................................................................... 35

Chapter 3. Conceptual Framework ................................................................................. 37

3.1 Definitions of the Key Concepts ................................................................................... 37
  3.1.1 Social Interaction, Social Relations, Friendship, and Network Ties ......................... 37
  3.1.2 Structure and the network setting ............................................................................. 39
  3.1.3 Activity of users ....................................................................................................... 40

3.2 Trajectory of Tie Growth at the Individual Level ......................................................... 41
  3.2.1 Variability of online individual friendship growth .................................................. 41
  3.2.2 Uniformity of online individual friendship formation ............................................. 43

3.3 Trajectory of Tie Growth at the Global Level .............................................................. 45

3.4. Explanatory Model of Friendship Formation .............................................................. 47
  3.4.1 Tie formation: Back to a general theory of social interaction ................................. 47
  3.4.2 Factors influencing online friendship formation ...................................................... 50

Chapter 4. Research Methodology .................................................................................. 68

4.1 Data Collection ............................................................................................................. 68

4.2 Measurement ................................................................................................................ 69
  4.2.1 Self-disclosure ......................................................................................................... 69
  4.2.2 Users popularity ..................................................................................................... 70
  4.2.3 Balance Structure .................................................................................................. 70
  4.2.4 Cohesiveness of user’s ego network ....................................................................... 71
  4.2.5 Experience and homophily of experience .............................................................. 72
  4.2.6 Other homophily variables ..................................................................................... 72

4.3 Anthropological Data Mining ..................................................................................... 73
Abstract

The online social networking sites (SNSs), which have rocketed from technique engaging niche users into a dominant online communication platform engaging millions of users, may change the pattern of online interpersonal communication. The study found that dyadic characteristics and structural factors largely determined the behavior of friendship formation, rather than characteristics of users.

The data of the current study are collected from one of the largest SNS platforms in China, coving the first two years (2006-07) since its inception, during which time most users of the SNS were college students. The dataset records an anonym zed ID, demographic characteristics, a list of bidirectional friends and the time of the friendship formation. In total, 10 million users and 200 million bidirectional ties are involved.

Using polynomial logistic regression to fit the time path of friendship formation for each individual user, the study finds that the increase in the number of friends for each user typically follows a Logistic function with time, indicating that the growth of friends starts slowly at the beginning, speeds up rapidly after reaching a critical point, and then tapers off finally. More importantly, the trajectories appear uniformly, if not identically, across individuals who joined the social network at quite different points in time, suggesting a high degree of regularity in user behavior when forming friendship.

This study further examines the growth trajectory of friendship of the entire SNS, and compares the resulting of global-level trajectory with the individual-level trajectories. Surprisingly, there emerges a strong self-similarity in the growth of friendship between the
individual level and the global level, which reinforces the regularity of social friendship formation behavior.

Although users show a similar time path of adding friends, there are substantial differences across individuals in the sheer size of their friendship networks. As the second research objective, the study examines the effects of three sets of factors, including characteristics of network structure, characteristics of dyadic friendship, and characteristics of individual users, on the likelihood of the friendship formation.

Of the factors under study, Balance Structure and the Cohesiveness of User’s Ego Network (which involves the user as the “ego-center” and only directly-related friends as other members of the network) are the two most dominant factors influencing the tie formation. Contrary to the prevailing theory of “Preferential Attachment” in the previous literature, friendship formation is positively related to Homophily of Degree (i.e, number of friends), which means that people make friends with similar (rather than famous) others. In addition, the Popularity of the user and his/her Experience with using the SNS has a negative effect on friendship formation, i.e., users with higher popularity or longer SNS using history are less likely to add new friends. Likewise, Homophily of Experience has also a negative effect on friendship formation in that users become increasingly inclined to establish friendship with those with different duration of experience with the SNS (i.e., veteran users are more likely to add new comers whereas new comers more likely to team up with veterans), which suggests that friendship formation is a dynamic process throughout which users change their friendship strategies over time. Finally, individual characteristics, such as Sex and level of online Self-
Disclosure, do not have any significant effect on the tie formation, which in fact strengthens the importance of structural factors reported above.

The findings of structural effects on friendship formation sheds some lights on social media marketing in the sense that online marketers could choose marketing information initiators according to users network positions, rather than users SNS usage frequency or motivation. In addition, this study also has some practical implications on SNS operators. Considering the strategy of retaining user loyalty, SNS operators should focused more on the friendship maintenance, such as encouraging the creation of user generated contents and developing tools to facilitate information sharing, rather than simply encouraging friendship formation since most users will stop adding new friends after they reach their saturation point.
List of Tables

Table 1. Summary of Friendship Formation Studies .......................................................... 136
Table 2. Summary of Models for Network Evolution and Dynamics ................................... 138
Table 3. Turner’s General Social Interaction Theory .......................................................... 139
Table 4. Three Types of Users .......................................................................................... 140
Table 5. Correlation Matrix (Significance Level) ............................................................... 141
Table 6. Assessing the PH Assumption .......................................................................... 143
Table 7. Model Estimation ............................................................................................... 144
List of Figures

Figure 1. Dynamics of studies on Online and Offline Friendship Formation ........................................ 145
Figure 2. Frequency of Theories used in Friendship Formation Studies .................................................. 146
Figure 3. Distribution of Methods used in Friendship Formation Studies ................................................. 148
Figure 4. Social Interaction and Social Relations ......................................................................................... 147
Figure 5. Demonstration of Dijkstra’s single source shortest path algorithm ........................................... 148
Figure 6. Distribution of R-squares of PLR modeling ................................................................................. 150
Figure 7. Life Span of 24 Cohorts ............................................................................................................... 151
Figure 8A. Distribution of Saturation Points (percent) ............................................................................. 153
Figure 8B. Distribution of Saturation Point (day) .................................................................................... 154
Figure 9. Model Comparison of the Simulation Data ................................................................................ 155
Figure 10A. Network Growth of Node and Edges (Population Network) .................................................. 156
Figure 10B. Network Growth of Node and Edges (Sampled Network) .................................................... 157
Figure 11. Evolution of Mean Degree (Sampled Network) ..................................................................... 158
Figure 12. Degree Distribution of the Sub-Network ................................................................................ 159