

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Yue Zhang

Entitled

A CROSS-SITE STUDY OF USER BEHAVIOR AND PRIVACY PERCEPTION IN SOCIAL NETWORKS

For the degree of Master of Science

Is approved by the final examining committee:

Melissa Dark

Brandeis H. Marshall

Baijian Yang

To the best of my knowledge and as understood by the student in the *Thesis/Dissertation Agreement, Publication Delay, and Certification/Disclaimer (Graduate School Form 32)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Melissa Dark

Approved by Major Professor(s): _____

Approved by: Gene Spafford

07/08/2014

Head of the Department Graduate Program

Date

PREVIEW

A CROSS-SITE STUDY OF USER BEHAVIOR AND PRIVACY PERCEPTION IN
SOCIAL NETWORKS

A Thesis

Submitted to the Faculty

of

Purdue University

by

Yue Zhang

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

August 2014

Purdue University

West Lafayette, Indiana

UMI Number: 1582809

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 1582809

Published by ProQuest LLC (2015). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

To my Mom, Dad and Qi, for all the support and love they give me.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my committee. I am so fortunate to have the best committee to help me going through the process of inquiry. My Chair, Prof. Dark has been patiently giving me advice and leading me to the right direction. I would never forget her support, advice and encouragement. I want to thank Prof. Yang and Prof. Marshall for their unwavering help and invaluable input. I also want to thank Prof. Whitten for helping me going through the formatting process.

There are many who have supported the creation of this work, making it difficult to choose the words that truly express the heartfelt gratitude and appreciation. I feel I am fortunate to be surrounded by an amazing group of graduate students and faculties. Thank every one of them and thank Purdue for giving me such a wonderful experience.

TABLE OF CONTENTS

	Page
LIST OF TABLES	vii
LIST OF FIGURES	xi
LIST OF ABBREVIATIONS	xiv
GLOSSARY	xv
ABSTRACT	xvi
CHAPTER 1. INTRODUCTION	1
1.1 Motivation and Objectives	2
1.2 Significance	3
1.3 Research Questions	4
1.4 Assumption	5
1.5 Limitations	5
1.6 Delimitations	6
1.7 Summary	7
CHAPTER 2. REVIEW OF LITERATURE	8
2.1 Types of Social Networks	8
2.2 Information Disclosure Behavior across Multiple Social Networks	9
2.3 Personal Traits and Privacy Perceptions	17

	Page
2.4 Summary	18
CHAPTER 3. METHODOLOGY	20
3.1 Research Population and Sample Size	20
3.2 Data Collection	22
3.3 Data Analysis Methods	23
3.4 IRB Protocol	23
3.5 Summary	23
CHAPTER 4. PRESENTATION AND ANALYSIS OF DATA	24
4.1 Data Summary	24
4.1.1 General Privacy Attitude	27
4.1.2 Privacy Setting	29
4.1.3 Privacy Policy	30
4.1.4 Extent of Trust	31
4.1.5 Profile Preference	33
4.1.6 Constituents of Friend List	34
4.1.7 Information Disclosure	36
4.2 Exploring Cultural Differences	37
4.2.1 Cultural Differences in General Privacy Attitude	38
4.2.2 Cultural Differences in Trust	44
4.2.3 Cultural Differences in Reading Privacy Policies	51
4.2.4 Cultural Differences in Privacy Settings	56
4.2.5 Cultural Differences in Friend Lists	62

	Page
4.2.6 Cultural Differences in Profile Preferences	67
4.2.7 Cultural Differences in Information Disclosure.....	72
4.3 Privacy Attitude and Information Disclosure Behavior.....	78
4.4 Relationships among Privacy Perceptions and Behaviors	93
4.5 Summary	104
CHAPTER 5. DISCUSSION OF RESULTS	105
5.1 Modeling Privacy Attitude, Perception and Behavior	106
5.2 Cultural Differences in Privacy Attitude	109
5.3 Cultural Differences in Privacy Perceptions and Behaviors.....	115
5.4 Privacy Attitude, Perception and Behavior in Different Sites	122
5.5 Conclusion and Future Work	130
LIST OF REFERENCES.....	133
APPENDICES	
Appendix A Survey Questions.....	135
Appendix B IRB Protocol	144

LIST OF TABLES

Table	Page
Table 2.1 Information Disclosure in Different OSNs (Schrammel et al. 2009)	11
Table 2.2. Personal Information Disclosed in Social Networks (Irani et al. 2009)	13
Table 3.1. Estimated Population and Targeted Sample Size	21
Table 4.1. Dataset Summary	24
Table 4.2. Number of Users for Each Site	26
Table 4.3. General Privacy Attitude	28
Table 4.4. Frequency of Changing Privacy Settings.....	29
Table 4.5. Have You Read Privacy Policies?	31
Table 4.6. Extent of Trust	32
Table 4.7. Profile Preference	33
Table 4.8. Constituents of Friend List	35
Table 4.9. Information Disclosure	37
Table 4.10. Results for Testing Differences in General Privacy Attitudes.....	42
Table 4.11. Constituents of Respondents for Each Social Network.....	45
Table 4.12. Extent of Trust: U.S. Citizen	47
Table 4.13. Extent of Trust: Chinese in the U.S.	47
Table 4.14. Extent of trust: Chinese in China.....	48
Table 4.15. Summary of Results for Extent of Trust.....	49

Table	Page
Table 4.16. Reading Privacy Policies: U.S. Citizen	53
Table 4.17. Reading Privacy Policies: Chinese in the U.S.	53
Table 4.18. Reading Privacy Policies: Chinese in China.....	54
Table 4.19. Summary of Results for Reading Privacy Policies	55
Table 4.20. Changing Privacy Settings: U.S.....	58
Table 4.21. Changing Privacy Settings: Chinese in the U.S.....	59
Table 4.22. Changing Privacy Settings: Chinese in China	59
Table 4.23. Summary of Results for Changing Privacy Settings	60
Table 4.24. Percentage of Real World Friends in Friend List: U.S. Citizen	63
Table 4.25. Percentage of Real World Friends in Friend List: Chinese in the U.S	64
Table 4.26. Percentage of Real World Friends in Friend List: Chinese in China.....	64
Table 4.27. Summary of Results for Friend List	65
Table 4.28. Profile Preference: U.S. Citizen.....	68
Table 4.29. Profile Preference: Chinese in the U.S.	68
Table 4.30. Profile Preference: Chinese in China.....	69
Table 4.31. Summary of Results for Profile Preference	70
Table 4.32. Information Disclosure: U.S. Citizen	73
Table 4.33. Information Disclosure: Chinese in the U.S.	74
Table 4.34. Information Disclosure: Chinese in China.....	75
Table 4.35. Cultural Differences in Information Disclosure	76
Table 4.36. Privacy Attitude vs Information Disclosure: Facebook.....	79
Table 4.37. Privacy Attitude vs Information Disclosure: Twitter	82

Table	Page
Table 4.38. Privacy Attitude vs Information Disclosure: WhatsApp	84
Table 4.39. Privacy Attitude vs. Information Disclosure: RenRen	86
Table 4.40. Privacy Attitude vs. Information Disclosure: Weibo.....	89
Table 4.41. Privacy Attitude vs. Information Disclosure: WeChat	91
Table 4.42. Correlation Matrix for All Facebook Users	93
Table 4.43. Correlation Matrix U.S. Facebook Users.....	94
Table 4.44. Correlation Matrix All Twitter Users	95
Table 4.45. Correlation Matrix U.S. Twitter Users	95
Table 4.46. Correlation Matrix Chinese Twitter Users in the U.S.	96
Table 4.47. Correlation Matrix All WhatsApp Users.....	97
Table 4.48. Correlation Matrix U.S. WhatsApp Users	97
Table 4.49. Correlation Matrix Chinese WhatsApp Users in the U.S.	98
Table 4.50. Correlation Matrix All RenRen Users	99
Table 4.51. Correlation Matrix Chinese RenRen Users in the U.S.	99
Table 4.52. Correlation Matrix Chinese RenRen Users in China.....	100
Table 4.53. Correlation Matrix All Weibo Users	100
Table 4.54. Correlation Matrix Chinese Weibo Users in the U.S.....	101
Table 4.55. Correlation Matrix Chinese Weibo Users in China	101
Table 4.56. Correlation Matrix All WeChat Users	102
Table 4.57. Correlation Matrix Chinese WeChat Users in the U.S.	103
Table 4.58. Correlation Matrix Chinese WeChat Users in China.....	103
Table 5.1. Rankings of Privacy for Different Types of Information	111

Table	Page
Table 5.2. Privacy Indices for Different Cultural Contexts	111
Table 5.3. Privacy Scores for Different Cultural Contexts	118
Table 5.4. Privacy Indices for Different Cultural Sites	111
Table 5.5. Privacy Scores for Different Sites	118

PREVIEW

LIST OF FIGURES

Figure	Page
Figure 2.1 Faceted Identities (Farnham et al. 2011)	15
Figure 4.1 How Many Sites Do you Use?	25
Figure 4.2 Number of Users for Each Site.....	26
Figure 4.3 General Privacy Attitude	27
Figure 4.4 Frequency of Changing Privacy Settings	29
Figure 4.5 Have You Read Privacy Policies?	30
Figure 4.6 Extent of Trust.....	32
Figure 4.7 Profile Preference	33
Figure 4.8 Constituents of Friend List	35
Figure 4.9 Information Disclosure.....	36
Figure.4.10 General Privacy Attitudes.....	39
Figure.4.11 One-way ANOVA Results	40
Figure.4.12 Post-hoc Results	41
Figure.4.13 Extent of Trust: U.S. citizen.....	45
Figure.4.14 Extent of Trust: Chinese in the U.S.....	46
Figure.4.15 Extent of Trust: Chinese in China	46
Figure.4.16 Kruskal-Wallis Test Results: Trust	49
Figure.4.17 Reading Privacy Policies: U.S. Citizen	51

Figure	Page
Figure.4.18 Reading Privacy Policies: Chinese in the U.S.....	52
Figure.4.19 Reading Privacy Policies: Chinese in China	52
Figure.4.20 Kruskal-Wallis Test Results: Privacy Policy	54
Figure.4.21 Changing Privacy Settings: U.S. Citizen.....	57
Figure.4.22 Changing Privacy Settings: Chinese in the U.S.	57
Figure.4.23 Changing Privacy Settings: Chinese in China.....	58
Figure.4.24 Kruskal-Wallis Test Results: Privacy Setting	60
Figure.4.25 Percentage of Real World Friends in Friend List: U.S. Citizen	62
Figure.4.26 Percentage of Real World Friends in Friend List: Chinese in the U.S.....	62
Figure.4.27 Percentage of Real World Friends in Friend List: Chinese in China	63
Figure.4.28 Kruskal-Wallis Test Results: Friend List	65
Figure.4.29 Profile Preference: U.S. Citizen	67
Figure.4.30 Profile Preference: Chinese in the U.S.	67
Figure.4.31 Profile Preference: Chinese in China	68
Figure.4.32 Kruskal-Wallis Test Results: Profile Preference.....	70
Figure.4.33 Information Disclosure: U.S. Citizen	72
Figure.4.34 Information Disclosure: Chinese in the U.S.....	72
Figure.4.35 Information Disclosure: Chinese in China	73
Figure 4.36 Privacy Attitude vs Information Disclosure: Facebook	79
Figure.4.37 Privacy Attitude vs Information Disclosure: Twitter	81
Figure.4.38 Privacy Attitude vs Information Disclosure: WhatsApp.....	84
Figure.4.39 Privacy Attitude vs Information Disclosure: RenRen	86

Figure	Page
Figure 4.40 Privacy Attitude vs. Information Disclosure: Weibo	88
Figure.4.41 Privacy Attitude vs. Information Disclosure: WeChat.....	91
Figure 5.1 The Cube Model	107
Figure 5.2 The Cube Model-Construction Phase.....	108
Figure 5.3 The Cube Model-Usage Phase	109
Figure 5.4 Privacy Attitudes vs. Actual Information Disclosure – Facebook	114
Figure 5.5 Radar Chart: Privacy Scores for Different Cultural Contexts	118
Figure 5.6 Radar Chart: Privacy Indexes for Different Cultural Contexts	119
Figure 5.7 Radar Chart: Privacy Scores for Different Sites.....	123
Figure 5.8 Radar Chart: Privacy Indexes for Different Sites 1	124
Figure 5.9 Radar Chart: Privacy Indexes for Different Sites 2	126
Figure 5.10 Radar Chart: Privacy Indexes for Different Sites 3	127
Appendix Figure	
Figure B.1 Original IRB Protocol	144
Figure B.2 IRB Approval of Amendment.....	145

LIST OF ABBREVIATIONS

OSN - Online Social Networks

SSN - Social Security Number

ANOVA - Analysis of variance

URL - Uniform Resource Locator

IRB - Institutional Review Board

PREVIEW

GLOSSARY

Privacy attitude	User's general privacy concern toward information privacy. To be specific, it means what information a user considers private and what information she considers public.
Privacy perception	Measures how a user perceives privacy risks while using social networks. In this thesis, such perception refers to how much a user trusts social networking sites.
Privacy behavior	A user's actual behavior that relates to privacy protection or indicates privacy awareness while using social networks. Such behavior includes changing privacy setting, using private profile, etc.
One-way ANOVA	A common technique used to compare means of two or more samples. It tests the null hypothesis that samples in two or more groups are drawn from populations with the same mean values.
Kruskal-Wallis test	Non-parametric equivalent of one-way ANOVA. It's used for testing differences of ordinal variables in this study.
Cohen's guideline	A guideline for interpreting correlation results. According to this guideline, $r = 0.5$, 0.3 and 0.1 represents large/medium/small correlation respectively (Cohen 1988).

ABSTRACT

Zhang, Yue. M.S., Purdue University, August 2014. A Cross-site Study of User Behavior and Privacy Perception in Social Networks. Major Professor: Melissa Dark.

While online social networking sites have brought convenience and diversity in people's social lives, they have also been the source for information leakage. Researchers have been looking for ways to balance user privacy protection and information disclosure. However, literature suggested that many users either failed to perceive privacy risks correctly or they failed to behave in accordance with privacy awareness even they have already perceived potential risks.

This thesis conducted a survey to measure social network users' privacy attitude, privacy perception and their actual behavior when using social networking sites. The survey targeted at three populations of different cultural contexts: U.S. college students, Chinese students in the U.S. and Chinese students in China. It also targeted at 6 popular sites – Facebook, Twitter, WhatsApp, RenRen, Weibo and WeChat.

Based on the survey results, this thesis conducted a cross-cultural and cross-site study to explore the relationships of social network users' privacy attitudes, privacy perceptions and various user behaviors. It also studied whether cultural contexts and the differences of sites had an impact on privacy attitude, perception and behavior.

CHAPTER 1. INTRODUCTION

The Internet's wide adoption has contributed to online social networking sites' thriving popularity. Facebook, for example had 1.3 billion monthly active users worldwide in 2014 (StatisticBrain, 2014) compared with a total of 835 million in 2012 (Internet World Status, 2012). The nature of social networks is to imitate real world social relationships by providing mechanisms for sharing information, creating personal profiles, establishing relations and communicating with each other. Not surprisingly, at the same time of using such services, people are giving out massive amount of information which may pose real threat to privacy. Documented threats include identity theft, digital stalking, and personalized spam. The problem becomes worse when most people are completely unaware of short-term and long term risks of sharing personal information without restricted access (Schrammel et al., 2009; Krishnamurthy et al., 2008; Acquisti et al., 2006) .

Privacy preserving methods which aim at anonymizing the social graph (Machanavajjhala et al., 2007; Li et al., 2007; Sweeney, 2002), privacy setting management (Squicciarini et al, 2012) or raising privacy awareness by evaluating user's privacy score (Liu et al, 2010), etc. have not been proven successful in protecting user's sensitive information or changing user's information disclosure behavior.

This thesis argued that in order to preserve privacy in social networks, the differences and interactions of multiple social networks should be considered as well as the differences of user's privacy attitudes to better define and mitigate privacy risks. This study took the first step towards evaluating and preserving privacy by studying the differences and interactions of user's privacy, attitude, perception and behavior variables in different social networking sites. It also studied whether people of different cultural contexts would perceive or behave differently when using social networks.

1.1 Motivation and Objectives

Recent literature has identified that using multiple social networks have become an emerging threat to user privacy. The study by Irani et al. (2011) has shown that the more social network a user uses, the more information can be potentially leaked. They argued that, because different social networks have different privacy protections, the risk of information leakage may be dependent on the "weakest point" in the social network ecosystem. Malhora et al. (2012) successfully linked the different profiles in different social networks that belonged to the same user which demonstrated the threats for those users who used multiple social network services.

Therefore, it would no longer be valid that privacy protection can be contained within the boundary of each social network. The information flow among social networks enables profit-seeking individuals or organizations to collect as much "digital footprints" (Irani et al. 2011) as possible by integrating a user's information that she has disclosed from all the sites that she uses.

Though such an emerging risk has been raised for several years, few literatures have focused on such topic and no solutions that attempt to preserve privacy across multiple social networks that have been proven effective.

The above literature assumed that the privacy risks came from the fact that users disclosed different information in different social networks. However, arguments such as “users do not behave inconsistently nor they have inconsistent profiles in different social networks” or “users may not care about the information they have provided at all” may easily debunk the above assumption. To find out a solution that adapts to real life scenario, the fundamental understanding of why and how the usage of different social networks poses threat to privacy is necessary.

Therefore, the main objective of this study is to understand how and why user’s privacy perceptions and behaviors differ in different social networks and among different cultural contexts and how the privacy perception influences the information disclosure behavior.

1.2 Significance

As discussed above, previous studies have assumed that users behave differently in social networking sites, however, none of them sufficiently justified their argument either because of lacking empirical data to support their argument or they fail to explore the reasons behind the differences of information disclosure.

Schrammel et al. (2012) took a site-centered approach to explore the difference of information disclosure on different types of social networks. This aggregated approach failed to distinguish the difference within the same type of social networks. Wang’s work

on the other hand, did discover the difference of user's tagging behavior between two popular bookmarking websites. However, this research only focuses on only two tagging sites and such an ad-hoc result can't be justifiably generalized to other sites.

To the best of our knowledge, a cross-site study of privacy attitude, perception and behavior on different social networks has not yet been conducted studied the impact of cultural contexts on social network users which have been rarely documented.

Another significance of the study was that it provided an up-to-date survey that investigates diverse aspects of social network privacy. Hopefully it could help researchers in this field better understand the usages, perceptions, attitudes and behaviors of social network users.

1.3 Research Questions

The questions central to this research are as follows:

1. What are users' privacy attitudes when they use the social networks? (e.g. what information do they consider private and what is not?) Are they different?
2. Does culture background have a significant impact on social network users' privacy perceptions and behaviors?
3. What's the relationship among a user's privacy attitudes, perceptions, and behavior in a specific site?

1.4 Assumption

The following assumptions are inherent to this study:

1. The participants are assumed to be honest and to have a basic understanding of the definition of privacy and social networks without major confusion.
2. The participants are assumed to be able to use the Qualtrics online survey system and to navigate and answer the questions correctly.
3. The participants will not retake the survey as not to disproportionally affect the outcomes.
4. The participants are representative of the study population.
5. The survey provides adequate information for the research questions.

1.5 Limitations

The study has the following limitations:

1. The self-reported survey may be biased as the actual behavior of the respondents on social networks may be different from what they reported. The self-reported survey may be biased also because the non-respondents may be more concerned about privacy.
2. Social networking is fast-evolving. The popularity of each site rises and falls. The results of this study including the survey itself are prone to be dated.
3. The design of the survey questions tried to capture the general characteristics of each sites. In many cases, they were of coarse-grained. The profile preference question for example, only had two options – public or private