May 2018

Three Essays on Trust Mining in Online Social Networks

Gelareh Towhidi
University of Wisconsin-Milwaukee

THREE ESSAYS ON

TRUST MINING IN ONLINE SOCIAL NETWORKS

by

Gelareh Towhidi

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctoral of Philosophy

in Management Science

at

The University of Wisconsin-Milwaukee

May 2018

ABSTRACT

THREE ESSAYS ON TRUST MINING IN ONLINE SOCIAL NETWORKS

by

Gelareh Towhidi

The University of Wisconsin-Milwaukee, 2018
Under the Supervision of Professor Atish P. Sinha and Professor Huimin Zhao

This dissertation research consists of three essays on studying trust in online social networks. Trust plays a critical role in online social relationships, because of the high levels of risk and uncertainty involved. Guided by relevant social science and computational graph theories, I develop conceptual and predictive models to gain insights into trusting behaviors in online social relationships.

In the first essay, I propose a conceptual model of trust formation in online social networks. This is the first study that integrates the existing graph-based view of trust formation in social networks with socio-psychological theories of trust to provide a richer understanding of trusting behaviors in online social networks. I introduce new behavioral antecedents of trusting behaviors and redefine and integrate existing graph-based concepts to develop the proposed conceptual model. The empirical findings indicate that both socio-psychological and graph-based trust-related factors should be considered in studying trust formation in online social networks.

In the second essay, I propose a theory-based predictive model to predict trust and distrust links in online social networks. Previous trust prediction models used limited network structural data to predict future trust/distrust relationships, ignoring the underlying behavioral trust-inducing factors. I identify a comprehensive set of behavioral and structural predictors of trust/distrust links based on related theories, and then build multiple supervised classification models to predict trust/distrust links in online social networks. The empirical results confirm the superior fit and predictive performance of the proposed model over the baselines.

In the third essay, I propose a lexicon-based text mining model to mine trust related user-generated content (UGC). This is the first theory-based text mining model to examine important factors in online trusting decisions from UGC. I build domain-specific trustworthiness lexicons for online social networks based on related behavioral foundations and text mining techniques. Next, I propose a lexicon-based text mining model that automatically extracts and classifies trustworthiness characteristics from trust reviews. The empirical evaluations show the superior performance of the proposed text mining system over the baselines.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and appreciation to my dissertation co-chairs, Dr. Atish Sinha and Dr. Huimin Zhao, for their guidance, help, and encouragement over the last two and a half years. I am deeply grateful for their continuous support, valuable advice, and their sincere interest in my success. I would like to thank Dr. Mark Srite for his helpful insights and feedback, especially in helping me develop and test the conceptual model. I am also very thankful to Dr. Sanjoy Ghose for his constructive criticism, insightful suggestions, and advice, particularly from a marketing standpoint. My extended thanks go to the Lubar School of Business for providing me with the research fellowships, project assistantships, and teaching opportunities, which were an invaluable part of my academic experience. I would especially like to thank my amazing family members for their endless love, support, and encouragement. I am deeply grateful to my husband for taking care of our son and providing me with emotional support. I am also grateful for my parents' financial and emotional support, and their constant encouragement to pursue my doctoral degree.

CHAPTER 1

INTRODUCTION

The rapid growth of online social networks in recent years has brought considerable attention to social network analysis. Recent studies have examined the structures and dynamics of social networks and their actors' relationships and interactions. Trust plays a critical role in online social interactions, because of the high levels of uncertainty and risk envolved. The important role of trust in e-commerce interactions and virtual teams has been studied before. However, trusting behaviors in online social networks __ how trust forms and what factors influence trusting behaviors __ have not been well studied. Furthermore, the existing models of trust prediction to predict trust and distrust relationships do not have a strong theoretical foundation. Hence, the established behavioral theories of trust need to be either adapted to the context of online social interactions or new ones need to be developed.

This dissertation introduces new theory and advances techniques for social network analysis by developing conceptual and predictive models of trust in online social networks. It consists of three research essays to study the issue. The first essay proposes a conceptual model of trust formation in online social networks to understand online trusting behaviors. The second essay develops a theory-based predictive model to predict trust and distrust links in online social networks. The third essay proposes a lexicon-based text mining model to mine user-generated content (UGC), and extract and classify trustworthiness characteristics from trust reviews. The findings of this dissertation would be of potential benefit to a wide range of applications, such as recommender systems, advanced search engines, online marketing, and social network trend prediction.

Essay 1: Trust Formation in Online Social Networks – A Network-Based Socio-Psychological Model of Online Trusting Behaviors

The first essay studies trust formation in online social networks. Trust plays a more critical role in online social relationships, compared to face-to-face communications, because of the higher levels of risk and uncertainty involved. However, trusting behaviors in online social networks have not been well studied yet. Previous research on trust in online social networks has been mostly conducted in the field of computational network analytics, ignoring the underlying theoretical dimensions of trusting behaviors. Guided by relevant psychology, social science, and computational graph theories, I propose a conceptual model of trust formation in online social networks. This is the first study that proposes a theory-based conceptual model to understand trusting behaviors in online social networks. The model integrates the existing graph-based view of trust formation in social networks into socio-psychological theories of trust. To develop the proposed model, I introduce new behavioral antecedents of trusting behaviors based on related socio-psychological theories. Next, I redefine and integrate the existing graph-based structural factors of online social networks into the model. I also study the joint influence of behavioral trust-inducing factors and graph-based factors on trusting behavior. The behavioral trust-inducing factors are measured according to the context of online social networks, using novel operationalization methods. The empirical findings indicate that both socio-psychological and graph-based trust-related factors should be considered in studying trust formation in online social networks. In particular, trust formation in online social networks is influenced by online trusting beliefs, network-based dispositional trust, network-based judgment bias, network-based structural assurance, trustee network status, network community similarity, and network status differential.

Essay2: Trust Prediction in Online Social Networks – A Theory-Based Predictive Model for Trust/Distrust Prediction in Online Social Networks

The second essay focuses on predicting trust/distrust links between the members of online social networks. Link prediction, i.e., predicting if there will be a link or a relationship between a pair of users, is one of the core problems in social network analysis. In signed social networks, the sign (trust/distrust, friend/foe, etc.) of a link has to be determined in addition to the likelihood of the relationship. Previous research mostly considered trust prediction as a special case of sign prediction, focusing on computational models of trust, based on the network structural data. They have largely ignored the underlying theory-based behavioral predictors of trust. Most previous models predict future trust links based on very limited dimensions of trusting behavior similarity between users. Moreover, previous studies predict only trust links, ignoring the critical information from distrust relationships. Guided by relevant socio-psychological and social graph theories, I identify influential factors in online trusting behaviors and propose a supervised model to predict trust/distrust links in online social networks. The proposed trust/distrust prediction model is built upon both structural and behavioral predictors, and by using both trust and distrust information. Specifically, I propose trustee importance, trustee prestige, trustor trust propensity, trustor status, similarity between the trustee and trustor, and status differential between the trustee and trustor as the prediction feature sets. The empirical results from my evaluation show the superior fit and predictive performance of the proposed model, compared to previous models that used only limited structural information.

Essay 3: Lexicon-Based Trust Mining in Online Social Networks – A Feature-Based Text Mining Model to Mine Trust Reviews in Online Social Networks

The third essay explores the important trustworthiness characteristics in trusting decisions by mining trust-related UGC. This is the first study that examines trust-inducing factors in online trusting decisions by mining UGC. I first identify the important trustworthiness characteristics of online users based on the related behavioral theories. Based on those characteristics, I generate domain-specific trustworthiness lexicons for online social networks, using text mining techniques. The trustworthiness lexicons are built using both dictionary-based and corpus-based methods. Using the generated lexicons, I propose a lexicon-based text mining system based on term co-occurrence and Pointwise Mutual Information to extract and classify trustworthiness characteristics from trust reviews. The empirical results show the superior performance of the proposed text mining system compared to other baseline models.

CHAPTER 2

Essay 1: Trust Formation in Online Social Networks – A Network-Based Socio-Psychological Model of Online Trusting Behaviors

2.1. Introduction

The rapid growth of online social networks in recent years has brought considerable attention to social network analysis. *Link prediction* in social networks, i.e., predicting whether there will be a relationship (link) between two users (nodes) of a social network, is one of the core problems in social network analysis (Backstrom and Leskovec 2011; Fang et al. 2013; Liben-Nowell and Kleinberg 2007; Wasserman and Faust 1994). Understanding how and why links form between users can benefit a wide range of applications, such as recommender systems, advanced search engines, online marketing, and social network trend prediction. In many real social networks, e.g., Epinions, Slashdot, and Wikipedia, the sign of a connection (positive/negative, trust/distrust, friend/foe, etc.) is an intrinsic part of the connection. In such signed networks, in addition to predicting link formation, predicting the sign of a potential link (i.e., *sign prediction*) is also of great importance (Leskovec et al. 2010a). Previous research on trust in online social networks often considered the *trust prediction* problem as a special case of sign prediction (e.g., Liu et al. 2008; Matsuo and Yamamoto 2009; Tang et al. 2013; Zhang et al. 2013).

Trust is a critical component of any interpersonal and social relationships in which risk, uncertainty, or interdependence exists (Fukuyama 1995; Gefen et al. 2003). In online interactions, trust plays a more critical role because of the inherent higher level of uncertainty, risk, and fear of opportunistic behaviors. It has been argued that mediation by technology in technology-mediated communications creates challenges for effective social interactions. For example, there are social cue deficiencies in computer-mediated interactions since body language and physical surroundings cannot be easily realized through computer channels (Ma and Agarwal 2007; Sproull and Kiesler 1986).

Online communities are characterized by a number of special characteristics that make effective social interactions through establishment of mutual understanding even more difficult (Chidambaram and Tung 2005; Ma and Agarwal 2007). In online communities, there are a large number of participants with different social backgrounds and perspectives (Ma and Agarwal 2007). In addition, most users are anonymous or only limited information about them is publicly available.

Past research has shown the important role of trust in virtual teams and e-commerce (Gefen et al. 2003; Hoffman et al. 1999; McKnight et al. 2002). However, trust in online social networks—how trust forms and what factors influence trusting behaviors—has not been well studied. Related studies were mostly in the field of computational network analytics, using existing knowledge and methods from sign prediction to examine trust formation.

Trust computational models used existing users' trust voting (e.g., "web of trust") to predict future trust relationships. Users' similarity, interactions, social status, global network structural features, local node properties, and network balance were identified as important factors in trust formation in online social networks (Chiang et al. 2011; Kunegis et al. 2009; Leskovec et al. 2010a; Leskovec et al. 2010b; Symeonidis et al. 2010; Varlamis et al. 2010). However, the underlying psychological and behavioral antecedents of trusting behaviors have been widely ignored, resulting in an incomplete understanding of users' trusting behaviors in online social networks.

Trusting behaviors of online users in social media and online communities differ from the traditional face-to-face and e-commerce settings. There are characteristics specific to an online social network that make it distinct from other settings. This requires traditional antecedents of trusting behaviors, as well as online social network-specific antecedents, to be examined in context.

The purpose of relationships in online social networks is obviously different from e-commerce buyer-seller relationships. The main purpose of social media and online social networks is to support interpersonal communication and collaboration using Internet-based platforms (Kane et al. 2014). Furthermore, social media and online social networks have provided new capabilities and opportunities that are not available in traditional online or offline settings. Kane et al. (2014) argue that one of the challenges in studying online social networks or social media is behavioral. They believe that the technological distinctions by social media may result in profound theoretical consequences for individual and organizational behavior.

Social media has provided its users novel ways of acting and interacting with each other that would have been difficult or impossible in earlier online or offline settings. The new capabilities of social media and online social networks may undermine or violate the assumptions of established theories (Kane et al. 2014). Hence, researchers need to adapt established theories for application to social media settings, or develop new ones (Kane et al. 2014; Majchrzak 2009). Therefore, I believe that there is a strong need to examine the applicability of related behavioral theories of trust in the context of online social networks.

There are characteristics unique to online social networks that may influence trusting behaviors. Garton et al. (1997) believe that there are three specific characteristics – range, centrality, and role – of online social networks, making them different from traditional work groups and communities. Online social networks are usually large. They have more heterogeneity in the social characteristics of network members and more complexity in the structure (Wellman 1997). In addition, centrality and connectedness of members of online social networks play a critical role in many aspects of sustainability of online social networks, such as information dissemination. Furthermore, it has been argued that there are network roles that may be filled by any member of a network according to what resources they bring in to the network, such as "technological gatekeeper". These roles can be identified by empirically examining the patterns of relations across networks or across behaviors within a network (Garton et al. 1997). Other researchers believe that social media differs from traditional social networks in terms of content and structure (Kane et al. 2014). Content refers to resources available in a network (such as information), and structure refers to identifiable patterns of nodes and ties in a network (Kane et al. 2014). For example, contagion and resource access are different from traditional networks (Borgatti and Foster 2003). Contagion is how resources spread through a network and influence nodes, and resource access is how nodes access and benefit from available resources. For example, in traditional groups, people were influenced by a very few numbers of their close friends, thus spread of an idea was usually locally. Therefore, I examine trusting behaviors in online social networks by examining network-specific trust inducing factors, in addition to traditional socio-psychological antecedents of trust.

In this study, I strive to address the gaps in the existing literature to explain trusting behaviors in online social networks. I propose a novel conceptual model of trust formation guided by relevant behavioral and computational network theories. To the best of my knowledge, this is the first study to provide a theory-based conceptual model of trust formation in online social networks. I introduce new behavioral constructs that impact online trusting behaviors based on related socio-psychological theories that have not been studied before. Moreover, I redefine and integrate existing graph-based factors along with the theory-based behavioral factors to proposed the model of trust formation in online social networks. I empirically examine the proposed conceptual model using a real social network dataset concerning the Wikipedia election process for RfA (Request for Adminship). The results show that the previous graph-based view of trust in social networks is not sufficient to understand trust formation among social network users, and both socio-psychological and graph-based trust-related factors should be considered.

2.2. Related Work

Online trusting behaviors have been studied mainly in the context of e-commerce and buyer-seller relationships, showing the important role of trust in online purchase intentions of consumers (e.g., Awad and Ragowsky 2008; Ba and Pavlou 2002; Everard and Galletta 2005; Fang et al. 2014; Gefen et al. 2003; Gefen and Straub 2003; Harris and Goode 2010; Kim and Benbasat 2006; Kim and Benbasat 2009; Komiak and Benbasat 2006; Lim et al. 2006; Liu and Goodhue 2012; Lowry et al. 2008; McKnight et al. 2002; Wang and Benbasat 2008). McKnight et al. (2002) proposed a multidimensional model of trust in e-commerce, including four high-level constructs, i.e., disposition to trust, institution-based trust, trusting beliefs, and trusting intentions. Trust in e-commerce relationships was also studied with regard to some related theoretical models, such as the *Technology Acceptance Model (TAM)*. Gefen and Straub (2003) examined the effect of social presence on consumer trust in e-commerce and purchase intentions. They found that consumer trust, affected by social presence, has a stronger effect on purchase intentions than TAM beliefs (usefulness and ease of use of the website). Gefen et al. (2003) showed specifically that both trust in the e-vendor and the technological aspects of the website interface (two beliefs identified by TAM, i.e., usefulness and ease of use) influence consumer intended use.

Some scholars examined the role of feedback mechanism or WOM (Word of Mouth) in trust formation in e-commerce relationships. Ba and Pavlou (2002) examined trust formation in electronic markets and showed that appropriate feedback mechanisms can induce trust between two transacting parties. Lim et al. (2006) examined the effect of trust on actual online buying behavior and found that satisfied customer endorsement by similar peers increased consumers' trusting beliefs about the e-vendor and ultimately actual buying behaviors. Awad and Ragowsky (2008) examined the effect of WOM quality on online trust and adoption of e-commerce, along with other factors, such as perceived ease of use and perceived usefulness, across genders. Trust in recommendation agents (RAs), which provide product recommendations to buy based on user-specified needs and preferences, has also been studied. Komiak and Benbasat (2006) showed that perceived personalization and familiarity significantly increased customers' intention to adopt RAs by increasing cognitive trust and emotional trust. Wang and Benbasat (2008) identified six reasons users trust (or do not trust) the RAs. Finally, some scholars examined the effects of website design related factors, such as website trust-assuring arguments (Kim and Benbasat 2006; Kim and Benbasat 2009) and website quality (Lowry et al. 2008), on online consumers' trust in e-vendors. However, trusting behaviors among users of online social networks or online communities have received much less attention, as online social networks and communities are themselves new phenomena in online communications.

Trust prediction in online social networks is a special case of the sign prediction problem. Trust prediction aims to use existing knowledge and methods of sign prediction to predict trust/distrust formation between network users. Furthermore, trust prediction studies aim to examine the special characteristics of trust networks and important factors of trust/distrust formation. Trust prediction models in online social networks usually use users' previous trusting behaviors (e.g., "web of trust") to predict future trust relationships. Network structural features, as well as some contextual features, have been used to build trust prediction models. Previous studies found that similarity, interaction, structural features, and node properties are important factors in trust formation. Guha et al. (2004) developed a trust propagation framework to predict trust between a pair of users and showed that distrust had a significant effect on how trust propagates through the network. Liu et al. (2008) proposed a supervised prediction model to predict trust links between a pair of users based on user attributes and user interactions in an online community. Matsuo and Yamamoto (2009) used supervised prediction models to predict trust based on users' profiles, product ratings, and trust relations. Tang et al. (2013) proposed an unsupervised framework (hTrust) for trust prediction, specifically focusing on the effect of homophily on trust prediction. Their findings showed that similar users tend to establish trust relations and that trusted users are more similar. Zhang et al. (2013) proposed a method to predict trust in networks by combining local information, such as node degree, clustering coefficient, and PageRank values, with the concepts of longer cycles and triangle patterns. Previous studies on trust prediction, however, have mostly focused only on the structural properties and dynamics of networks and did not consider any underlying theory-based behavioral motives of trust formation.

2.3. Theoretical Background and Hypothesis Development

Trust is a multi-dimensional concept and has been studied in multiple contexts. Trust is defined as a trustor's willingness to be vulnerable to the actions of a trustee based on positive expectations of the trustee's intentions (Mayer et al. 1995; Mcknight et al. 1998; Rousseau et al. 1998). Trust has been conceptualized according to several theoretical streams, including knowledge-based trust, calculative-based trust, institution-based trust, cognition-based trust, and personality-based trust (Gefen et al. 2003; Mayer et al. 1995; Mcknight et al. 1998).

Trusting Behavior in Online Social Networks

Almost all definitions of trust mention "willingness to rely on the other party or to be vulnerable" based on a "positive expectation or beliefs in the other party's trustworthiness" (Mayer et al. 1995; Mcknight et al. 1998; Moorman et al. 1992; Rousseau et al. 1998). Based on this, trust has a cognitive aspect (*trusting beliefs* or beliefs in the other party's trustworthiness) and a behavioral aspect (*trusting intentions* or willingness to rely on another party) (Mcknight et al. 1998; Moorman et al. 1992). These two aspects have been studied as two distinct constructs, in which trusting beliefs affect trusting intentions and together they represent trusting in another party (Mcknight et al. 1998; McKnight et al. 2004; Moorman et al. 1992).

Behavioral intention has been shown as an antecedent of actual behavior, as modeled in many theories, such as the *Theory of Reasoned Action* (Fishbein 1979; Fishbein and Ajzen 1975) and the *Theory of Planned Behavior* (Ajzen 1985; Ajzen 1991; Ajzen and Madden 1986). Although the intention-behavior relationship has been reliably shown in prior studies (Webb and Sheeran

2006), several variables can influence the consistency of this relationship (Sutton 1998; Webb and Sheeran 2006). Thus, if possible, measuring the actual behavior is more reliable than measuring behavioral intentions. In the context of online social networks, in contrast to most behavioral studies where the actual behavior cannot be measured, access to users' actual trust decisions is feasible. Therefore, in this study, the actual trusting decision is used as trusting behavior. *Trusting behavior* could be one of the following possible links (trust relationships): a trustor trusts a trustee, a trustor distrusts a trustee, or a trustor is neutral toward a trustee. For example, in some online review websites, users can build a web of trust composed of reviewers whose reviews they trust or block reviewers whose reviews they distrust. When user *a* adds user *b* to her web of trust, it means user *a* (the trustor) issues a trust link toward user *b* (the trustee). When user *a* adds user *b* to her block list, it means user *a* (the trustor) issues a distrust link toward user *b* (trustee).

Online Trusting Beliefs

Beliefs about a behavior have been viewed as important antecedents of behavioral intentions and behaviors in many theories. The Theory of Reasoned Action (TRA) (Fishbein 1979; Fishbein and Ajzen 1975) states that behavioral intentions, the immediate antecedent to behavior, are influenced by beliefs about the outcome of the behavior, including behavioral beliefs and normative beliefs. Behavioral beliefs affect an individual's attitude toward the behavior and normative beliefs affect the individual's subjective norm about the behavior (Madden et al. 1992). The Theory of Planned Behavior (TPB) (Ajzen 1985; Ajzen 1991; Ajzen and Madden 1986) adds beliefs about control over performing the behavior, which focuses on beliefs about possessing the required resources and opportunities, to the TRA.

Trust involves taking risk and selecting a trustworthy partner could eliminate some of the concerns in trust relationships (Sheppard and Sherman 1998). Based on that, trustworthiness is a quality that mitigates the concerns in trust relationships, such as ability, reliability, honesty, and altruism. *Trusting beliefs* (also called interpersonal trust) is defined as a positive sentiment or expectation about the trustee's trustworthiness (Mcknight et al. 1998; Moorman et al. 1992). Multiple scholars have studied trusting beliefs and identified four main dimensions of trusting beliefs, i.e., *benevolence*, *competence*, *integrity* (Butler 1991; Gabarro 1978; Mayer et al. 1995; Mcknight et al. 1998), and *predictability* (McKnight et al. 2002; Mcknight et al. 1998). Benevolence is the trustor's belief that the trustee acts in the trustor's best interest and cares about the trustor. Competence is the trustor's belief in the trustee's ability or power to do what is needed by the trustor. Integrity is the trustor's belief in the trustee's truthfulness, that she makes good faith agreements, tells the truth, and fulfills promises. Predictability is the trustor's belief that the trustee's actions (good or bad) are consistent enough to forecast them. These four dimensions are conceptually distinct (Kumar et al. 1995), but have been used in combination as a global measure of trusting beliefs (Doney and Cannon 1997; McKnight et al. 2002; McKnight et al. 2004). A trustor's intention to trust a trustee is related to the extent of the trustor's beliefs in the trustee's benevolence, competence, integrity, and predictability (Mcknight et al. 1998).

Prior research in trust has shown the links between trusting beliefs and trusting intentions and modeled trusting beliefs as the immediate antecedent of trusting intentions (Mayer and Davis 1999; Mayer et al. 1995; Mcknight et al. 1998). Furthermore, the relationship between trusting beliefs and trusting intentions has also been supported in the field of e-commerce and buyer-seller relationships (Gefen 2002; McKnight et al. 2002). Thus, I expect trusting beliefs to be related to trusting behaviors in online social communities too. For example, if Emily thinks that

John is an expert in movies and he always tweets high-quality and honest reviews about new movies, then it is very likely that Emily trusts John and follows him on Twitter. Hence, I hypothesize the following:

H1: Trusting beliefs (benevolence, competence, integrity, and predictability) toward a trustee are positively related to trusting the trustee.

Network-Based Dispositional Trust

Dispositional trust (i.e., a trustor's *trust propensity*) is the trustor's inherent tendency to generally trust others across various situations (Mayer et al. 1995; Mcknight et al. 1998). Trust propensity is an individual’s personal characteristic that determines how much the individual is willing to trust others in general. A trustor’s trust propensity is based on her faith in humanity and trusting stance (Mcknight et al. 1998). Faith in humanity is one’s belief that others are typically well-meaning and reliable. It has been argued that since faith in humanity refers to underlying beliefs that in general others are trustworthy, it influences one’s trusting beliefs (Kramer and Isen 1994; Mcknight et al. 1998). Moreover, disposition to trust is a generalized tendency across situations and people and is stable in different situations. As an inherent characteristic, it influences one’s interpretation of situations and actors in those situations. Thus, in trust relationships, it influences one’s interpretation of a trustee’s trustworthiness, providing a foundation for the trustor’s trusting beliefs (Gefen 2000; Mcknight et al. 1998). Trusting stance is the willingness to depend on others (whether they are trustworthy or not) and the belief that by dealing with people, better outcomes will be achieved. It has been argued that since trusting stance is a personal strategy in dealing with others, it directly influences one's willingness to trust others (McKnight et al. 2002; Mcknight et al. 1998).

Trustor’s trust propensity has been modeled as one of the important antecedents of trusting beliefs and trusting intentions (Mayer et al. 1995; McKnight and Chervany 2001; Mcknight et al. 1998; Williams 2001). Previous studies also found significant effect of dispositional trust (propensity to trust) on consumers’ trust in e-commerce relationships (Gefen 2000; McKnight et al. 2004). Disposition to trust is not based on prior experience or knowledge about a specific trustee. Thus, it is more important in initial trust formation when the parties have not gained specific knowledge or experience of each other.

In the context of online relationships that are distant and where limited information is available about the parties, disposition to trust could play an important role in initial trust formation (McKnight et al. 2004). This effect can be even more important in online social networks, where there is even less information about users than would be available about an e-commerce vendor. For example, Emily has a higher trust propensity than John, meaning that she believes more than John that others are typically well-meaning and trustworthy. Alex is a new user to Facebook and is recommended to both to become friends. Emily is more likely to have more positive beliefs in Alex’s trustworthiness than John. Consequently, she is also more likely to become a friend of Alex. Hence, I hypothesize the following:

H2a: Dispositional trust is positively related to trusting beliefs.

H2b: Dispositional trust is positively related to trusting the trustee.

Network-Based Judgment Bias

Psychological research in decision making suggests that human intuitive judgments deviate from a normatively expected standard of judgment subject to a number of biases, resulting in *biased judgment* (Funder 1987; Griffin and Tversky 1992; Tversky and Kahneman 1975). Bias in judgment happens because of cognitive processes, such as conservatism (Fiedler 1991), exaggerated expectation (Erev et al. 1994), placement judgments (Kruger 1999), and level of confidence in judgment (Tversky and Kahneman 1975). Judgment bias is almost a consistent systematic bias that could be used for predicting individual behavior (Hilbert 2012).

In the context of e-commerce relationships, Wolf and Muhanna (2011) found that online buyers systematically interpret online feedback rating information in a biased fashion, resulting in inaccurate trusting beliefs. This bias in processing feedback information results in inaccurate estimates of online retailers' trustworthiness, which can happen in two ways, *overconfidence* or positive bias (giving unwarranted trust in an e-vendor) and *underconfidence* or negative bias (being overly suspicious). Illusion in cognitive processing could also result in biased judgments, such as overconfidence in judgment (Langer 1975). The logic is that people in uncertain situations try to assure themselves that things are under control, resulting in unrealistic overly positive perceptions that differ from reality (Fiske and Taylor 1984). Trust research also suggests that trust building involves illusions of control process in a way that the cognitive process enhances one's confidence in trust-related beliefs, resulting in overconfidence bias (Mcknight et al. 1998). Mcknight et al. (1998) argued that illusions of control process happen in building a trustor's trusting beliefs. The trustor tries to confirm her trusting beliefs in the other party's trustworthiness, thereby overinflating her confidence that high levels of trusting beliefs are warranted.

Trust prediction in online social network research also posits that the truthfulness (bias) of users' judgments about others should be considered in predicting trust (Mishra and Bhattacharya 2011). This is important because, in many social networks, the prestige of a user is based on the opinion or rating of other users. The prestige of a node can be used to predict the trustworthiness of the node or estimate how much a node is trusted by other users. If a user gives positive (or negative) ratings to others irrespective of what they truly deserve, then the user is biased. The truthfulness (or bias) of a user's judgment is the difference between the trust rating a trustor provides to a trustee and the ground truth (i.e., what the trustee truly deserves). Therefore, trust ratings from biased users should be given less weight in trustworthiness prediction of users in social networks.

According to the above theoretical and empirical foundations, judgment bias is very likely to happen in evaluating the trustworthiness of users in trust decisions. Thus, a trustor's judgment about a trustee's trustworthiness is very likely to be affected by the trustor's judgment bias. A trustor with overconfidence bias has overinflated confidence that high levels of trusting beliefs are warranted. Thus, there is a positive relationship between judgment bias and trusting beliefs in others, no matter of what they truly deserve. Hence, I hypothesize the following:

H3: Judgment bias is positively related to trusting beliefs.

Cognitive biases are raised because of people's limited ability to properly process information. Judgment bias, as a cognitive bias, is the systematic errors in judgments that deviate systematically from an accepted norm or standard. One of the salient causes of judgment bias is preconceived ideas or theories about people and events (Kruglanski and Ajzen 1983). An individual's reliance on intuitive preconceptions and theories leads to systematic judgment biases

(Kruglanski and Ajzen 1983; Tversky and Kahneman 1975). The reason is that people tend to consider only certain information and disregard other relevant information based on their intuitive ideas about the event or behavior under consideration. I posit that this is true about people's disposition to trust and its effects on their judgment bias. Disposition to trust, as an inherent tendency to trust others, can act as a filter for receiving and processing information accurately. A person with high trust propensity is incapable of processing the information in an unbiased way, because she inherently selects positive evidence that others are trustworthy and tends to disregard any negative clues about them.

It has also been argued that there is an interaction between one's disposition to trust (trust propensity) and illusions of the control process, which result in judgment bias (Mcknight et al. 1998). Disposition to trust brings the illusions that everything is always under control and affects an individual's trusting judgments. Furthermore, previous empirical studies have shown that individual characteristics and psychological attributes influence an individual's judgment bias (Barber and Odean 2001; Chen et al. 2007; Wolf and Muhanna 2011). For example, online buyers show different levels of judgment bias and trust in online vendors, based on their perceived uncertainty regarding the general other and the transaction medium or environment (Wolf and Muhanna 2011).

Based on the discussed theoretical background and empirical findings, I posit that trust propensity affects judgment biases in trusting decisions. I further posit that this is especially true in online relationships, where less prior interaction experience or lower levels of information exist. As argued by Mcknight et al. (1998), disposition to trust is more important at the beginning of a relationship, when beliefs about the situation are based more on assumptions than on facts.

Therefore, in online relationships, where actors have less prior information and facts about one another, their intuitive preconceptions, such as their propensity to trust, can have a very important role in their trusting decisions. For example, Emily, who has a high trust propensity (meaning that she believes that others are typically well-meaning and trustworthy), as an editor of Wikipedia, is selected to vote for new trusted admins of the community. She reviews John's (one of the candidates) request for adminship, as well as other voters' discussions about his qualifications and his past activities on Wikipedia. There are almost equal number of positive reviews and negative reviews about John's trustworthiness for adminship. Since Emily has inherently high level of trust propensity, she is very likely to disregard the negative reviews and pay attention to positive reviews, resulting in a biased judgment and consequently trusting John. Hence, I hypothesize the following:

H4: Dispositional trust is positively related to judgment bias.

Network-Based Structural Assurance

Institution-based trust (also called impersonal trust), which refers to a trustor's beliefs that the necessary structures and favorable conditions are in place, facilitates situational success in a risky endeavor (Mcknight et al. 1998). Institution-based trust helps the trustor feel more comfortable to trust others in that specific situation and affects interpersonal trust (i.e., trusting beliefs). There are two types of institution-based trust: *situational normality* and *structural assurance* (McKnight et al. 1998). Situational normality is the sense that success is likely because the situation is normal or favorable (Baier 1986; Mcknight et al. 1998). In the context of this study, I assume that situational normality is the same for everyone since all users are acting in the same online community. Structural assurance is the sense that success is likely because protective structures, such as guarantees, contracts, regulations, promises, legal recourse, processes, or procedures, are in place. Structural assurance increases trust and affects trusting beliefs by giving the sense that parties in the situation are trustworthy (Gefen et al. 2003; Mcknight et al. 1998).

In the context of e-commerce, structural assurance is usually provided by a reputable third party that assures the e-vendor is trusted (Gefen et al. 2003). However, in online social networks, relationships among users are formed based on users' willingness to communicate and often no external party is involved. Network structural characteristics influence trust in social networks (Buskens 1998; Sherchan et al. 2013). A user's position in the network could be an indicator of the user's reputation. Reputation, as an aggregate of users' opinions about one another, is often measured by *centrality* measures (Katz 1953). Central users are the key users of the network that connect with more users and are the major channels of information dissemination (Wasserman and Faust 1994). Other structural measures have been used to measure a node's reputation in a network, such as *PageRank* (Page et al. 1999) and its variants, and *HITS authority* (Kleinberg et al. 1999). Such structural reputation measures have been used to predict influential and trusted users in online social networks (Song et al. 2007; Varlamis et al. 2010). Furthermore, Buskens (1998) found that users with a higher degree centrality (both in-degree and out-degree) have higher levels of trust. Tsai and Ghoshal (1998) showed that central actors are likely to be perceived as more trustworthy by other actors in the network. Previous studies have shown the importance of tie strength in perceived trustworthiness and trust relationships (Krackhardt et al. 2003; Levin and Cross 2004). Central users are more connected and have more interactions with others. Central users are major channels of information dissemination, making them known by many others, further resulting in stronger ties and greater levels of perceived trustworthiness and trust. Hence, I hypothesize the following:

H5a: Structural assurance is positively related to trusting beliefs.

H5b: Structural assurance is positively related to trusting the trustee.

Network-Based Status (Prestige)

Social *status* is one's standing in a social hierarchy as determined by respect, deference, or social influence (Ridgeway and Walker 1995). Social status represents one's value position as relative prestige, respect, honor, or deference in a structural arrangement (Berger et al. 1972; Thye 2000). Socio-psychological theories, such as the *Expectation States Theory* (Berger et al. 1982) and the *Status Characteristics Theory* (Berger et al. 1980; Wagner and Berger 1993) explain the emergence of status, beliefs about status, and its influence on group members' behaviors. People differ based on a set of social attributes, such as gender or a specific expertise, called status characteristics. Status characteristics shape beliefs about social worthiness, competence, and performance expectations, which in turn, shape the social interactions and the behaviors of individuals in a group (Berger et al. 1980; Wagner and Berger 1993). High-status actors of the group are believed to be more worthy, competent, and influential (Ridgeway 2001; Ridgeway and Walker 1995). Higher status is also associated with greater perceived trustworthiness and higher levels of trust (Cook et al. 2009).

It has been empirically shown that the relative status of an individual in a group is positively related to perceived trustworthiness qualities, such as the ability and integrity of the individual (Campos-Castillo and Ewoodzie 2014). In the context of online social networks, social status is defined as the position or rank of a user in the network and represents the degree of honor or prestige attached to that position. A node's status (prestige) in a network has been considered as an important predictor of received positive links (friendship or trust) from other nodes of the network (Kunegis et al. 2009). A user's prestige (status) in online social networks is often identified by incorporating both trust and distrust links toward the user, such as the Fans Minus Freaks (FMF) (Kunegis et al. 2009) measure, where fans are ones who trust the user and freaks are ones who distrust the user. Trusted users have a high number of fans and distrusted users a high number of freaks. Therefore, a trustee with a higher status is perceived to be more trustworthy and receives more trust links from other members of the network. Hence, I hypothesize the following:

H6a: Trustee status is positively related to trusting beliefs.

H6b: Trustee status is positively related to trusting the trustee.

Furthermore, status differences have been considered as an influential factor in interpersonal perception and behavior (Ridgeway and Berger 1986). It has been argued that as status is "relative", the perceived trustworthiness of individuals is also a relationship-level attribute (Campos-Castillo and Ewoodzie 2014). It means that who is seen as trustworthy and the extent of perceived trustworthiness of any given individual are also contingent on the relational context. The *status differential* between two actors will determine the degree to which an actor perceives trustworthiness qualities in the other actor (Campos-Castillo and Ewoodzie 2014). Therefore, the status differential between a trustee and a trustor has an important role in shaping the trustor's beliefs about the trustee's trustworthiness. As the trustor's status relative to the trustee decreases (the status differential increases), the trustor's perceived trustworthiness of the trustee increases.

The size of the status differential between two actors is determined based on the number of status characteristics (such as specific expertise) and the magnitude of the difference between the values of a given status characteristic (Berger et al. 1972).

Moreover, the *Status Theory*, as one of the online social network theories, was developed for directed social networks to better explain the observed edge signs and provides insights into the underlying social mechanisms (Leskovec et al. 2010a; Leskovec et al. 2010b). Based on the Status Theory, a trust (distrust) link from a trustor to a trustee indicates that the trustee has a higher (lower) status than the trustor (Leskovec et al. 2010a; Leskovec et al. 2010b). This also reinforces the discussed relational characteristic of status by previous socio-psychological scholars. Therefore, the status differential between a trustor and a trustee is associated with greater perceived trustworthiness of the trustee (the trustor's trusting beliefs). Hence, I hypothesize the following:

H7: Trustee and trustor status differential is positively related to trusting beliefs.

It has been found that status influences behavioral levels of trust such that individuals with higher status show more initial trust than individuals with lower status (Lount and Pettit 2012). High social status brings the individual a set of internalized beliefs and expectations around the rewards she will receive by virtue of this elevated social position (Berger et al. 1998). The reward expectations involve social rewards, which induce positive expectations about others' motives and behaviors. Individuals with higher status have expectations that others will have favorable motives and display positive behaviors toward them, such as respect and praise. Thus, people with higher status judge others' intentions as more benevolent, and subsequently, trust others more easily (Berger et al. 1998; Lount and Pettit 2012). I posit that this could also be true for online trust interactions. Therefore, a trustor's status in the network could affect her beliefs and expectations about others. High-status trustors perceive others to be more benevolent and have positive expectations in others' motives and behaviors. Therefore, a trustor's status could affect her trusting beliefs toward a trustee. Hence, I hypothesize the following:

H8a: Trustor status is positively related to trusting beliefs.

A set of social attributes, such as gender or experience, which differentiate individuals' status in a group, are considered status characteristics. Previous research on behavioral decision making has shown that individual characteristics, such as gender, ethnicity, education, and job experience, affect an individual's judgment biases. For example, it has been found that self-serving attribution bias is greater for men than for women, resulting in men's overconfidence bias about their abilities (Barber and Odean 2001). Overconfidence and representativeness biases have been found in financial investors based on life experiences and the education system (Chen et al. 2007). Chinese investors were found to be more overconfident about their knowledge and skills, and believed more that past returns were indicative of future returns (a representativeness bias), than US investors. Based on these results, status characteristics can affect an individual's judgment biases.

Previous studies have also found that high-status individuals have more overconfidence bias in self-judgment than others (Harvey 1953; Stolte 1978). They tend to view their qualities in overly positive ways, thus inflating their self-perceptions and resulting in overconfidence about their abilities (Anderson and Brion 2010; Berger et al. 1980; Ridgeway et al. 1998). It has also been found that overconfident individuals attain higher social status in groups (Anderson et al. 2012).

Therefore, individuals with higher status are positively biased about themselves, meaning that they see their abilities higher than the reality.

Furthermore, it has been found that overconfident individuals think others are underconfident or unbiased, whereas underconfident individuals think others are overconfident (Ludwig and Nafziger 2011). People who believe that underconfidence is the general bias (who see others as being underconfident) adjust their beliefs upwards, thus resulting in being overconfident, and vice versa. Therefore, individuals with higher status are negatively biased about others, meaning that they see others' abilities lower than the reality.

Therefore, individuals' status plays an important role in their beliefs and judgment biases about themselves and others. Taking into account all these findings, I posit that individuals with higher status tend to have overconfidence bias about themselves and underconfidence bias about others and vice versa. Thus, I conjecture that a trustor with high (low) status tends to judge others' trustworthiness lower (higher) than their real trustworthiness. Hence, I hypothesize the following:

H8b: Trustor status is negatively related to the trustor's judgment bias.

Network-Based Similarity (Homophily)

Social science has identified *homophily* as an important factor in social networks and group formation. Early studies in network ties showed substantial homophily in demographic and psychological characteristics in informal network ties, such as friendship networks (McPherson et al. 2001). Homophily is the principle that people are more likely to bond and associate with similar others than with dissimilar ones, based on sociodemographic, behavioral, and intrapersonal characteristics (McPherson et al. 2001). Similarity among individuals affects their initial willingness to trust each other (Foddy et al. 2009). Similarity predisposes people toward a greater level of interpersonal attraction, understanding, and trust, resulting in greater levels of social affiliation (Brown et al. 2007; Ruef et al. 2003).

Homophily has also been observed and studied in online social network relationships (e.g., Wang et al. 2011; Xiang et al. 2010). Furthermore, homophily (i.e., similarity among users) has been used in link formation and sign (trust/distrust or friend/foe) prediction in social networks. Multiple studies have shown that similarity is an important predictor of trust formation among users in social networks (Golbeck 2009; Kunegis et al. 2009; Symeonidis et al. 2010; Tang et al. 2013; Zheng et al. 2014).

Homophily implies that people's personal networks are homogeneous and information flows through networks tend to be localized (McPherson et al. 2001). Therefore, there should be multiple localized communities of neighbors based on users' similarities in social networks. It has been found that having a shared group membership enhances trustworthiness perceptions, thereby increasing the levels of initial trust (Meyerson et al. 1996; Tanis and Postmes 2005). In the context of social networks, a shared group membership can translate into being in the same community with common neighbors. Having common neighbors is also found to be important in user collaboration in social networks. Newman (2001) found that the probability of collaboration increases with the number of other collaborators in common. Common links (neighbors) have also been widely used to measure the similarity between users in social networks, such as in the *Adamic-Adar index*, and in the *Jaccard Coefficient* (Adamic and Adar 2003; Golder and Yardi 2010; Liben-Nowell and Kleinberg 2007; Lichtenwalter et al. 2010). Accordingly, I consider common neighborhood as a similarity measure among users. Thus, a trustor's trust in a trustee increases as the trustor has more common neighbors with the trustee.

It has also been argued that users of the same group or community are more likely to form trusting beliefs toward one another (Mcknight et al. 1998). People in the same group share common goals and values, tend to perceive each other in a positive light (Kramer 1994), and are more trustworthy than others not in their group (Brewer and Silver 1978) (also known as ingroup bias). Therefore, having common neighbors leads to high levels of trusting beliefs. Thus, a trustor's trusting beliefs in a trustee increase as the trustor has more common neighbors with the trustee. Hence, I hypothesize the following:

H9a: Trustee and trustor community similarity is positively related to trusting beliefs.

H9b: Trustee and trustor community similarity is positively related to trusting the trustee.

2.4. The Research Model

The research model for testing the hypotheses stated above is shown in Figure 2.1. Definitions of the research constructs are presented in Table 2.1.

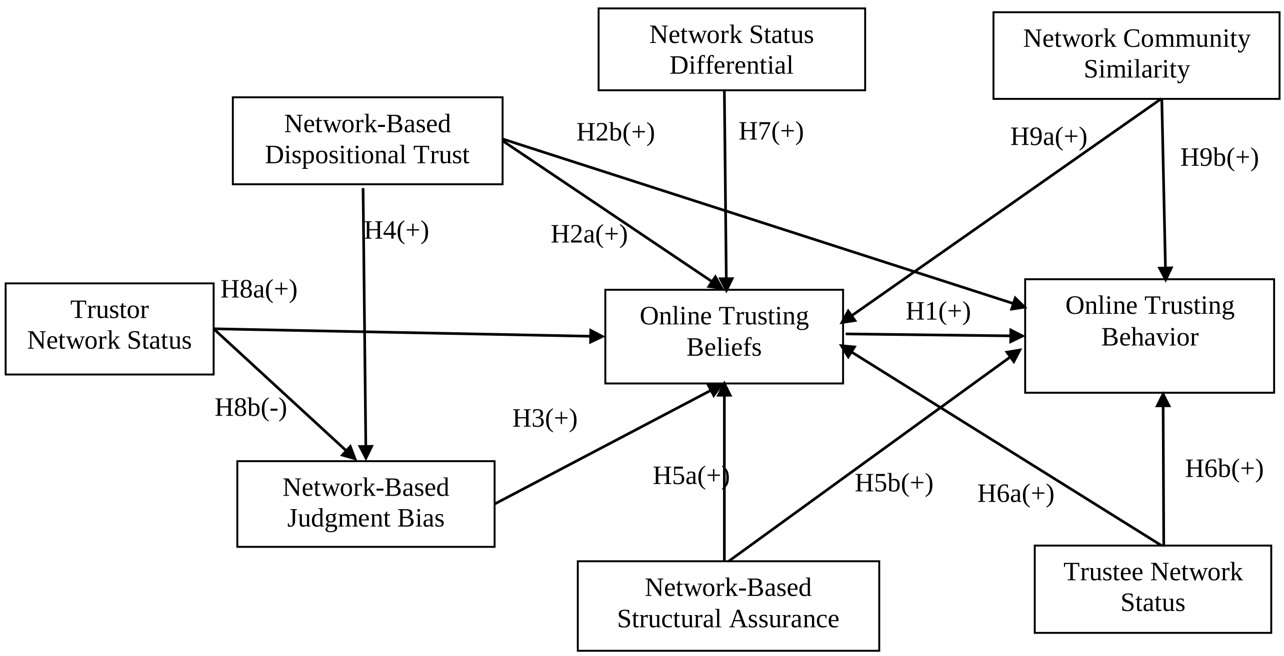

Figure 2.1- The Research Conceptual Model

Table 2.1- Research Constructs

Construct	Definition	References
Online Trusting Behavior	The decision to rely on the trustee or to be vulnerable by her actions in an online social network.	Mayer et al. 1995; McKnight et al. 1998; Rousseau et al. 1998
Online Trusting Beliefs	Beliefs in the trustee's trustworthiness according to the context of a particular online social network (benevolence, competence, integrity, and predictability).	Mayer et al.1995; McKnight et al. 1998
Network-Based Dispositional Trust	The inherent tendency to generally trust others across various situations based on the proportion of the trust out-degree to the total out-degree.	Mayer et al. 1995; McKnight et al. 1998
Network-Based Judgment Bias	Deviation from a normatively expected standard of judgment; the average bias of the trustor's total out-degree based on the trustee's true merit.	Funder 1987; McKnight et al. 1998
Network-Based Structural Assurance	The sense that the trustee is trustworthy because protective structures are in place; the centrality, HITS authority, and PageRank importance of the trustee node.	Gefen et al. 2003; McKnight et al. 1998
Trustor/Trustee Network Status	Node's standing or value position in an online social network; based on the net value of trust/distrust in-degree.	Berger et al. 1972; Thye 2000; Ridgeway and Walker 1995
Network Status Differential	The difference between trustee status and trustor status in the network.	Berger et al. 1972; Leskovec et al. 2010a; 2010b
Network Community Similarity	The degree that the trustor and the trustee share common neighbors in the network; Jaccard, Dice, and inverse log-weighted similarities.	McPherson et al. 2001; Newman 2001

2.5. Research Methodology

In the following section, the operationalization of research constructs, data set, and data analysis methodology are discussed.

2.5.1. Construct Operationalization

To empirically examine the proposed research model, I operationalized the constructs within the context of online social networks. The definitions for the constructs from the hypothesis development section of the paper are refined along with more detailed explanations of any sub constructs and terms. Herein I explain how each construct was measured.

Online Trusting Behavior

I used a 7-point Likert scale (“strongly oppose,” “oppose,” “weakly oppose,” “neutral,” “weakly support,” “support,” and “strongly support”) to measure the trusting decision of a trustor regarding a trustee. I used this measure as a single reflective indicator of the trusting behavior latent variable.

Online Trusting Beliefs

Trusting beliefs are formed based on the trustor’s cognitive process about the trustee’s trustworthiness in terms of the trustee’s benevolence, competence, integrity, and predictability (Mayer et al. 1995; Mcknight et al. 1998). A trustee is perceived to be benevolent to the degree of openness, loyalty, concern, and support and help and acting in the other party's best interests and caring about their well-being (Colquitt et al. 2007; Mayer and Davis 1999; Mayer et al. 1995; McKnight et al. 2002). Benevolence relies more on personal experience or interactions between the trustor and the trustee. A trustee is perceived to be competent as she has enough capability, proficiency, ability, expertise, knowledge, and talent in a specific context and performs her role very well and in an effective way (Colquitt et al. 2007; Mayer and Davis 1999; McKnight et al. 2002). A trustee's integrity is the extent of her truthfulness, honesty, fairness, promise keeping, commitment, sincerity and authenticity, bias suppression, ethicality of decision making, and credibility (Colquitt et al. 2007; Mayer and Davis 1999; Mayer et al. 1995; McKnight et al. 2002). Unlike benevolence, integrity is based more on the trustee's characteristics than on the interactions or relationships between the trustee and the trustor (McKnight and Chervany 2001). A trustee's predictability is the extent of her consistency of actions (Mcknight et al. 1998).

Based on the definitions and measurements discussed above, and the definitions and keywords of trustworthiness specific to the Wikipedia context, I created four sets of measurement items for the trustee's trustworthiness characteristics. Then I used the trustor's comments about the trustee and asked two graduate students to annotate the comments according to the measurement items. Each annotator identified if any of the trustworthiness characteristics are mentioned in the comments and the related sentiment. Then, I asked them to resolve the conflicts between their annotations, if any, to get the final values for all four trustworthiness beliefs.

The four dimensions (i.e., benevolence, competence, integrity, and predictability) of the trusting beliefs construct are formative in nature (Sia et al. 2009). All four dimensions represent distinct aspects of trusting beliefs and thus do not necessarily correlate highly. Therefore, I used each dimension as a separate formative indicator of the trusting beliefs latent variable.

Network-Based Dispositional Trust

I obtained trustor propensity in two different ways. First, I considered the proportion of the total trust links to the total issued links toward others: $Propensity(i) = \frac{outdegree^{+}(i)}{outdegree^{+}(i)+outdegree^{-}(i)}$, where $outdegree^{+}(i)$ is the out-degree of positive (trust) links of user i (the trustor) in the network, and $outdegree^{-}(i)$ is the out-degree of negative (distrust) links of the user. Second, I took into account both trust and distrust links: $Propensity(i) = \frac{outdegree^{+}(i)-outdegree^{-}(i)}{outdegree^{+}(i)+outdegree^{-}(i)}$. I used both measures as separate reflective indicators of the trust propensity latent variable.

Network-Based Judgment Bias

The bias of a trustor's judgment is the difference between the trust rating the trustor provides to a trustee and the ground truth (what the trustee truly deserves). I estimated the ground truth regarding a trustee based on the crowd wisdom (the final trusting decision from the community (all voters)). I obtained trustor's judgment bias in two different ways. First, I considered the difference between the trustor's vote to a trustee and the final trusting decision from the community for that specific trustee: $Bias(i) = \frac{(vote(i,j) - deserve(j))}{2}$, where $vote(i,j)$ is the vote of user *i* on user *j*, and $deserve(j)$ is the final voting result from the community for user *j*. The difference is divided by two to scale the bias value into the range of [-1, +1]. Second, I calculated the trustor's judgment bias using Mishra and Bhattacharya (2011) bias formula: $Bias(i) = \frac{1}{2|(d(i)|} \sum_{j \epsilon d(i)} (vote(i,j) - deserve(j))$, where $d(i)$ is the set of other users voted by user *i* (the trustor), $vote(i,j)$ is the vote of user *i* on user *j*, and $deserve(j)$ is the final voting result from the community for user *j*. The sum of the differences is normalized by the number of all the votes issued by the trustor and further divided by two to scale the bias value into the range of [-1, +1]. Therefore, a bias value of zero means that the trustor has no judgment bias (her votes are very close to what a trustee truly deserves). Positive (negative) values mean that the trustor is biased toward trusting others more (less) than what they really deserve. I used both measures as separate reflective indicators of the trustor's judgment bias latent variable.

Network-Based Structural Assurance

I consider the public view of a trustee as the general assurance of the trustee's trustworthiness. To operationalize structural assurance, I measured a trustee's general reputation in the network using multiple measures—centrality measures, HITS authority, and PageRank—by incorporating all the links (including trust, distrust, and neutral links) issued by all actors. By including all the links, I could examine how central and important the trustee is in the network.

Centrality measures (Freeman 1978) identify key nodes (users) for information dissemination in a social network, by providing leadership or bridging different communities (Chau and Xu 2012). *Closeness centrality* means nodes at the geographic center are central and measures how long it takes for information to pass from a trustee node to other nodes in the network. *Betweenness centrality* means nodes with many transits are central and measures how often a trustee node is found on the shortest path between two other nodes in the network. *Eigenvector centrality* means nodes connected to central nodes are central themselves. Hyperlink-Induced Topic Search (HITS) is a link analysis algorithm for rating web pages and consists of two measures: hub and authority (Kleinberg et al. 1999). A good hub represents a page that points to many other pages, and a good authority represents a page that is linked by many different hubs.

In other words, a hub is a node with many out-links and an authority is a node with many in-links. Here, I only consider the trustee authority, since a well-known trustee is one with many in-links rather than out-links. In the context of social graphs, PageRank is an iterative algorithm that measures the importance or authority of each node within the network. It has been discussed that a user (node) is influential if other influential users trust or follow her in a trust social network (Varlamis et al. 2010). I used Closeness centrality, Betweenness centrality, HITS Authority, and PageRank as separate reflective indicators of the structural assurance latent variable.

Network Status (Prestige)

Status is the aggregate opinion of others about a specific node in the network. A more prestigious user is one who is more trusted by others. I obtained trustee/trustor status by including the sign of the links issued to the user. In contrast to the structural assurance measures, in which we cared only about the centrality and intractability of users, here we focused on the type of the relationships. The simplest form of obtaining the status is through the *FMF (Fans Minus Freaks) measure* (Kunegis et al. 2009): $Status(i) = indegree^{+}(i) - indegree^{-}(i)$, where $indegree^{+}(i)$ is the in degree of positive (trust) links of user i in the network, and $indegree^{-}(i)$ is the in degree of negative (distrust) links of the user. I calculated the proportion of FMF to the total in-degree: $Status(i) = \frac{indegree^{+}(i) - indegree^{-}(i)}{indegree^{+(i)} + indegree^{-}(i)}$, and also to the total degree of the trustee node: $Status(i) = \frac{indegree^{+}(i) - indegree^{-}(i)}{degree\ (i)}$.

I used all of these three measures as separate reflective indicators of the trustor status latent variable.

I used two additional measures proposed by previous research to calculate trustee prestige. (Leskovec et al. 2010a; Leskovec et al. 2010b) introduced a new measure of trustee status in which a trustee node's status increases for each positive link it receives and each negative link it issues, and declines for each negative link it receives and each positive link it issues:

$$Status(i) = \frac{indegree^{+}(i) - indegree^{-}(i) - outdegree^{+}(i) + outdegree^{-}(i)}{degree(i)}$$

Furthermore, *Net Trust Votes (NTV)* is a measure of the Shapley value based centrality measures (Aadithya et al. 2010; Gangal et al. 2016) for directed signed networks. Since the main idea of the NTV is the same as the (Leskovec et al. 2010a; Leskovec et al. 2010b) status measure, I also considered it as a measure of trustee status: $Status(i) = \frac{1}{2}\left(indegree^{+}(i) - indegree^{-}(i)\right) - \frac{1}{2}\left(outdegree^{+}(i) - outdegree^{-}(i)\right)$

I used all the five status discussed measures as separate reflective indicators of the trustee prestige latent variable.

Status differential between a trustee and a trustor is calculated as the difference between the trustee's status and the trustor's status: *StatusDifferential* (*trustee, trustor*) = *Status*(*trustee*) – *Status*(*trustor*)

I used the four discussed status measures (except NTV) to calculate the status differential between a trustee and a trustor. Then, I used all the four status differential measures as separate reflective indicators of the status differential latent variable.

Network Community Similarity (Homophily)

Common friendships or common neighbors have been used widely to measure the similarity between two nodes. I used three commonly used similarity measures in social networks: Jaccard similarity, Dice similarity, and inverse log-weighted similarity (Adamic and Adar 2003). The *Jaccard similarity* coefficient is a commonly used similarity metric in information retrieval; it is defined as the number of matches of two sets divided by their union. Thus, it is the number of common neighbors divided by the number of nodes that are neighbors of at least one of the two nodes. The *Dice similarity* coefficient of two nodes is twice the number of common neighbors divided by the sum of the degrees of the nodes. The *inverse log-weighted similarity* of two nodes is the number of their common neighbors, weighted by the inverse logarithm of their degrees. I used all the three discussed similarity measures as separate reflective indicators of the similarity (homophily) latent variable.

2.5.2. Dataset and Data Analysis

To examine the proposed research model, I used Wikipedia adminship election data. Wikipedia administrators are the trusted members who are elected by the community members; they are granted additional tools to perform certain special actions. Members can request promotion from editor to admin status. Request for adminship (RfA) is an election process by which the community decides whether to trust or distrust the nominated candidates for adminship positions. The Wikipedia RfA election dataset has been specifically used as a real world example of signed online social networks in past studies (Chiang et al. 2011; DuBois et al. 2011; Leskovec et al. 2010a; Leskovec et al. 2010b). There are specific characteristics unique to Wikipedia RfA's authority-granting process that make the dataset particularly valuable to examine trusting behaviors among online community members. The structure of the network and interactions within the Wikipedia RfA authority-granting process creates a data set that clearly represent the characteristics of online social networks. The RfA is based on the election process in which members vote on the candidates' capabilities, meaning that they issue trust and distrust links toward other members. In the RfA process, Wikipedia members access other members' profiles and activities, judge their trustworthiness, and then issue relational connections. They can also view the connections (trust and distrust votes/links) made by other members, along with their opinions on the trustworthiness of those specific nodes. All voters express their reasons and opinions by posting comments, along with their votes (a 7-point Likert scale on the extent to which they trust the candidates). Thus, in addition to the fact that it is a real world sample of online trusting behaviors, members express their opinions about the trustworthiness of candidates. This enables us to measure trustor online trusting beliefs (benevolence, competence, integrity, predictability) about a trustee's trustworthiness. Another advantage is that all community members vote on the same set of candidates. That means we have many votes (trusting behavior/links) and opinions on those candidates. As a result, we do not face the issue of scarcity of links, which is a very common problem in online social networks. Assessing node importance or community similarity based on larger volumes of network data provides more realistic and accurate estimates of those measures. In addition, the RfA data set provides more reliable data of online trusting behaviors compared to other signed social networks. Members (voters) provide their justifications along with their votes to the community, based on the candidate's history and previous interactions, if any, with the candidate. Other voters can read those opinions and even discuss their viewpoints. Furthermore, community members vote for a shared goal, to select new admins, which is in the best interests of the entire community. Thus, trusting decisions are more reliable since members tend to be more careful and thorough. In contrast, trusting decisions in most other signed networks are just personal. The RfA dataset also provides unique information. For example, it has the final voting results (community decision) on candidates, which can be used for an in-depth analysis of trusting behaviors. The final decision from the community can be assumed to be a true merit of the candidates and can be used for measuring one of the important factors of trustors in trusting behaviors: judgment bias.

The data were collected from 2003 (since the adoption of the RfA process) through 2013, containing 11,402 users (voters and votees) and 198,275 votes (West et al. 2014) (also available from Stanford Network Analysis Project (SNAP)). All the network data, including all the users and all the issued votes (trust, neutral, distrust), were used to estimate the network related measures of the research model. I used *Structural Equation Modeling (SEM)*, specifically, the *Partial Least Squares (PLS)* method in SmartPLS 2.0, to analyze the data. PLS was used to test the psychometric properties of scales, measure reliability and validity (convergent and discriminant) of the constructs, and test the research model and hypotheses. PLS, as one of the SEM techniques, is a multivariate technique which allows estimating multiple equations simultaneously. PLS-SEM enables the measurement quality assessment via the measurement model and analyzing the relational effects via the structural path model. Furthermore, PLS-SEM can easily handle reflective and formative latent constructs in the same measurement model, as well as single-item constructs with no identified problem (Hair Jr et al. 2016). In addition, PLS-SEM has minimal restrictive assumptions compared to covariance-based SEM techniques.

I randomly sampled a balanced dataset (including all three types of votes) of 350 one-to-one relationships to test the research hypotheses. The “10-times” rule of thumb is widely used to estimate the required sample size for PLS-SEM (Hair Jr et al. 2016). The sample size for PLS should be equal to: either 10 times the largest number of formative indicators used to measure a single construct (if there are formative constructs in the model) or 10 times the largest number of structural paths to a construct in the structural model (largest structural equations (LSE) or arrows pointing to a latent variable). The largest number of structural paths to a construct (trusting beliefs) is seven and the largest number of formative indicators of a construct (trusting beliefs: benevolence, competence, integrity, predictability) is four. Taking into account both of them, the required sample size is 70 for PLS estimation. The sample size to test the research model is 350, which is more than adequate for PLS.

2.6. Results and Findings

In the following section, the measurement model validity and the structural model (hypothesis testing) results are discussed.

2.6.1. Measurement Model Validation

I evaluated the reliability, convergent validity, and discriminant validity of the constructs of the measurement model. PLS uses a set of iterative confirmatory factor analyses to validate the measurement quality (Hair Jr et al. 2016).

Reliability of measures was examined to assess if each scale consistently reflected the related construct it measured. Two measures, i.e., indicator reliability and internal consistency reliability (ICR), were used to assess the reliability of the measures. For indicator reliability, all outer loadings should be higher than 0.7 (Hulland 1999). Indicators with outer loadings less than 0.4 should be dropped and those between 0.4 and 0.7 should be considered for removal if the deletion increases the composite reliability and average variance extracted (AVE). AVE is used to measure convergent validity and is the degree to which a latent construct explains the variance of its indicators. Table 2.2 shows that all the outer loadings for constructs are above 0.7. The second reliability measure, ICR, weights items based on their factor loadings and is considered as a more robust reliability measure than Cronbach's alpha. ICR values should be higher than 0.7 (Bagozzi and Yi 1988); all values reported are above 0.7 (Table 2.2).

To evaluate the convergent and discriminant validities of the conceptual model, I used two measures (AVE and item loadings). For convergent validity, AVE values should be higher than 0.5 (Bagozzi and Yi 1988). For discriminant validity, the square root of average variance extracted by a construct from its indicators (AVE) should be greater than 0.7 and greater than the construct's correlation with other constructs (Fornell and Larcker 1981). Table 2.2 shows that all square roots of AVEs (number on diagonals) are above 0.7 (AVE>0.5) and greater than the loadings on other constructs, indicating that convergent and discriminant validities were verified.

Table 2.2- Measurement Quality Assessment for the Reflective Indicators

Construct	Items	Loading	ICR	AVE	TBL	JB	DT	TS	CS	SD	SA	TP	TBH
Online Trusting Beliefs (TBL)	Ben	Formative Measurement Model											
	Com												
	Int												
	Pre												
Network-Based Judgment Bias (JB)	JB1	0.772	0.779	0.638	0.278	0.799							
	JB2	0.824											
Network-Based Dispositional Trust (DT)	DT1	0.994	0.994	0.988	0.353	0.250	0.994						
	DT2	0.993											
Trustor Network Status (TS)	TS1	0.957	0.982	0.947	0.022	-0.058	0.022	0.973					
	TS2	0.976											
	TS3	0.985											
Network Community Similarity (CS)	CS1	0.985	0.967	0.908	0.129	0.006	0.088	0.359	0.953				
	CS2	0.881											
	CS3	0.988											
Network Status Differential (SD)	SD1	0.957	0.957	0.847	0.519	0.047	0.245	-0.484	-0.137	0.920			
	SD2	0.924											
	SD3	0.901											
	SD4	0.896											
Network-Based Structural Assurance (SA)	SA1	0.747	0.925	0.757	0.570	-0.169	0.184	0.035	0.206	0.492	0.870		
	SA2	0.943											
	SA3	0.825											
	SA4	0.949											
Trustee Network Prestige (TP)	TP1	0.98	0.965	0.825	0.589	-0.036	0.188	0.071	0.097	0.787	0.627	0.908	
	TP2	0.983											
	TP3	0.981											
	TP4	0.972											
	TP5	0.795											
Online Trusting Behavior (TBH)	TBH	1	1	1	0.936	0.303	0.377	0.031	0.159	0.554	0.588	0.633	Single-Item Construct
AVE: Average Variance Extracted; ICR: Internal Consistency Reliability													

All constructs were measured by items that are reflective indicators of their latent constructs, except for the trusting beliefs construct, which is formative in nature (Sia et al. 2009). The indicator items of the trusting beliefs construct represent distinct aspects of it and thus do not necessarily correlate highly. The indicators of trusting beliefs measure the trustor's beliefs about a trustee's trustworthiness qualities as benevolence, competence, integrity, and predictability. According to Hair Jr et al. (2016), the internal consistency reliability concept is inappropriate for formative indicators. Also, using similar criteria of reflective measurement models is not meaningful for assessing convergent and discriminant validities of formative indicators. The content validity assessment of formatively measured constructs should be done by assessing the convergent validity, potential collinearity issues, and the significance and relevance of the formative indicators.

To evaluate convergent validity, I tested whether the formatively measured trusting beliefs were highly correlated with a reflective measure of it (Hair Jr et al. 2016). I used the final decision of the trustor about a trustee's trustworthiness (i.e., trusting behavior) as a reflective measure of trusting beliefs. The strength of the path coefficient linking the two constructs was 0.93, which is above the desired magnitude (0.80) (Hair Jr et al. 2016). Furthermore, I did not find any collinearity problems among the indicators of the trusting beliefs. I ran a multiple regression analysis. All VIF (variance inflation factor) values were lower than 3 and all tolerance values were above 0.2, as shown in Table 2.3, indicating that there was no potential collinearity problem (Hair Jr et al. 2016).

Table 2.3- Collinearity Statistics of the Formative Indicators (Online Trusting Beliefs)

	Collinearity Statistics	
Indicator	**Tolerance**	**VIF**
Benevolence	0.816	1.226
Competence	0.731	1.368
Integrity	0.739	1.353
Predictability	0.707	1.415

Finally, the significance and relevance of the indicators of trusting beliefs were assessed. The outer weights of the indicators are the result of a multiple regression with the latent variable, representing the indicator's contribution (importance) to the construct. All outer weights were significant at the 0.001 level (Table 2.4), indicating that all are significant indicators of trusting beliefs. I therefore elected to keep all the four indicators (Hair Jr et al. 2016).

Table 2.4- Outer Weights for the Formative Indicators (Online Trusting Beliefs)

	Item Weight	t-stat
Benevolence → Online Trusting Beliefs	0.232	7.91 ***
Competence → Online Trusting Beliefs	0.456	13.70 ***
Integrity → Online Trusting Beliefs	0.403	12.41 ***
Predictability → Online Trusting Beliefs	0.281	8.89 ***
***: statistically significant at the 0.001 level.		

2.6.2. Structural Model Analysis (Hypothesis Testing Results)

To determine the explanatory power of the structural model and test the hypotheses, I examined the structural path coefficients (t-values) and the R-squared scores. PLS uses a bootstrapping procedure to test the statistical significance of the structural paths (Hair Jr et al. 2016). As suggested by Hair Jr et al. (2016), the sample size (350) was used as the number of "cases" and 5000 used as the number of "samples" in the bootstrapping process. The results of hypothesis testing are shown in Table 2.5. The model explains 89% (R^2 = 0.89) of the variance of online trusting behavior and 56% (R^2 = 0.56) of the variance of online trusting beliefs.

Trusting beliefs had a statistically significant positive effect on trusting behavior, as hypothesized in hypothesis H1. Trustor's propensity was positively related to trustor's trusting beliefs about the trustee and trusting behavior, supporting hypotheses H2a and H2b. Trustor's propensity to trust was also positively related to trustor's judgment bias toward trust, supporting hypothesis H4. The path from judgment bias to trusting beliefs was statistically significant, providing support for hypothesis H3. Structural assurance had statistically significant positive effects on both trustor's trusting beliefs and trusting behavior, as hypothesized in H5a and H5b, respectively. Trustee status also had a statistically significant effect on trusting behavior, supporting hypothesis H6b. However, it did not have statistically significant effect on trustor's trusting beliefs about the trustee; hypothesis H6a was therefore not supported. Trustee and trustor status differential was positively related to trusting beliefs, providing support for hypothesis H7. The effects of the trustor status on trusting beliefs and trustor's judgment bias were not statistically significant; hypotheses H8a and H8b were therefore not supported. In addition, the path between trustee and trustor community similarity and trusting beliefs was not statistically significant, hypothesis H9a was therefore not supported. However, its effect on trusting behavior was significant (hypothesis H9b).

I found that a trustor's trusting behavior toward a trustee is influenced by the trustor's trusting beliefs about the trustee's trustworthiness, the trustor's trust propensity and judgment bias, the structural assurance, the trustee's prestige, and their community similarity. A trustor's trusting beliefs about a trustee's trustworthiness characteristics (benevolence, competence, integrity, and predictability) had a statistically significant impact on trusting behavior. Trusting beliefs, in turn, were affected by the trustor's propensity to trust, trustor's judgment bias, structural assurance, and the status differential between the trustor and the trustee in the network. A trustor's trust propensity is an inherent characteristic of the trustor and is the trustor's general willingness to trust others. The more easily a trustor can trust others, the more positive beliefs she has about the trustworthiness of the trustee. A trustor's trust propensity had a statistically significant effect on her judgment bias toward the ground truth about a trustee's trustworthiness. A trustor with overconfidence judgment bias has overinflated confidence in her judgment about a trustee's trustworthiness, resulting in higher levels of trusting beliefs than the reality.

Table 2.5- Hypothesis Testing Results

Hypothesis	Path Coefficient	t-stat	Hypothesis Support
H1: Online Trusting Beliefs → Online Trusting Behavior	0.831	37.31***	Supported
H2a: Network-Based Dispositional Trust → Online Trusting Beliefs	0.116	2.88**	Supported
H2b: Network-Based Dispositional Trust → Online Trusting Behavior	0.054	6.78***	Supported
H3: Network-Based Judgment Bias → Online Trusting Beliefs	0.31	6.78***	Supported
H4: Network-Based Dispositional Trust → Network-Based Judgment Bias	0.251	4.76***	Supported
H5a: Network-Based Structural Assurance → Online Trusting Beliefs	0.393	9.07***	Supported
H5b: Network-Based Structural Assurance → Online Trusting Behavior	0.067	2.55**	Supported
H6a: Trustee Network Status → Online Trusting Beliefs	0.137	1.22	Not Supported
H6b: Trustee Network Status → Online Trusting Behavior	0.113	4.35***	Supported
H7: Network Status Differential → Online Trusting Beliefs	0.24	1.96*	Supported
H8a: Trustor Network Status → Online Trusting Beliefs	0.126	1.55	Not Supported
H8b: Trustor Network Status → Network-Based Judgment Bias	-0.063	1.19	Not Supported
H9a: Network Community Similarity → Online Trusting Beliefs	0.011	0.26	Not Supported
H9b: Network Community Similarity → Online Trusting Behavior	0.031	1.7*	Supported
*** Statistically significant at the 0.001 level ** Statistically significant at the 0.01 level * Statistically significant at the 0.05 level			

Although it has been theoretically argued that individuals with higher status have more positive trusting beliefs in others, I did not find any statistically significant support for this relationship. The path coefficient from the trustor's status to the trustor's trusting beliefs was positive (as hypothesized) but not statistically significant. This might have been because there are other influencing factors that contributed more toward explaining the variance of trusting beliefs. The personality-based factors (such as trust propensity and judgment bias), along with the relational (status differential) and situational (structural assurance) characteristics, had stronger effects on shaping the trustor's trusting beliefs than the trustor's status. The effect of the trustee's status on the trustor's trusting beliefs also was not statistically significant. Hence, the status of the trustor or the trustee in the group by itself was not an influential factor in shaping the trustor's trusting beliefs, rather it was their relational status that was important.

Previous research in the context of face-to-face relationships has also shown that higher-status individuals have an underconfidence judgment bias toward others. Although the path coefficient was negative, as hypothesized, the effect was not statistically significant. This might have been because of the stronger effect of the trustor's inherent propensity (dispositional trust) on the trustor's judgment bias. Since dispositional trust is a personality characteristic rooted in one's childhood, it was more determinant of the trustor's trusting beliefs than the trustor's status, which is not a permanent characteristic (i.e., it changes in different groups).

The general structural assurance was found to be a statistically significant antecedent of trusting beliefs and trusting behavior. Using structural reputation and centrality of the trustee, I showed that structural assurance was positively related to the trustor's trusting beliefs and trusting behavior toward the trustee. Generally, I found that more reputable and central users were assumed to be more trustworthy and were more trusted. Furthermore, the trustee's status (prestige) in the network also had a statistically significant positive effect on trusting behavior toward the trustee. The trustee's prestige takes into account the net trust values issued to the trustee.

I would like to note that prestige is different from centrality and general reputation, which were used to measure structural assurance. A central user is not necessarily a prestigious user, such that while the user might be connected with many others, her trust links are not necessarily more than her distrust or neutral links. However, the findings show that both centrality (reputability) and status (prestige) of the trustee were associated with greater trust from others in the network. While the trustee's status (prestige) statistically significantly influenced the trustor's trusting behavior, it did not statistically significantly affect the trustor's trusting beliefs. This might have been because of the effect of the crowd wisdom on individuals' decisions. A highly prestigious member is one who is trusted by the community, causing others to trust her anyway even if they do not deeply believe in her trustworthiness.

I also showed that the status differential between the trustor and the trustee was positively related to the trustor's trusting beliefs toward the trustee's trustworthiness. The lower the trustor's status than the trustee's status, the more positive the trustor's beliefs about the trustee's trustworthiness. Furthermore, the community similarity between the trustor and trustee (having common neighbors) was also statistically significantly related to trusting behavior. In line with previous findings, individuals tended to associate and bond with similar others and had more trust in them. While the community similarity between the trustor and trustee was not found to significantly affect the trustor's trusting beliefs, it directly affected the trustor's trusting behavior. This might have been because of the homophily principle, that the trustor automatically trusted similar others, even if she did not believe in their trustworthiness.

2.7. Discussion

In this research, I addressed trust formation among users of online social networks. I proposed a novel conceptual model for trusting behavior in online social networks and empirically examined its validity using real-world data.

In the context of online social networks, most of the previous studies focused only on the structural dynamics of social networks, ignoring the important behavioral factors in trusting behaviors. Furthermore, most of them focused only on a few structural characteristics of social networks to explain trust/distrust links, such as only similarity or node importance characteristics. Moreover, as argued by past researchers, there is a need to adapt traditional established theories for application to social media settings or possibly develop new ones (Kane et al. 2014; Majchrzak 2009) because of the capabilities of social media networks. The technological distinctions presented by social media may result in profound theoretical consequences for an individual's behavior (Kane et al. 2014). Hence, I was motivated to examine the validity of related behavioral theories in the context of online social networks. I strive to bring the social psychology research and computational network science together to explain trust formation in online social networks.

This study—the first study to provide a conceptual model of trust formation in online social networks—makes several major contributions. First, I introduce new behavioral constructs (antecedents) of trusting behaviors in online social networks. In particular, I identify and examine the effect of *online trusting beliefs*, *network-based dispositional trust*, and *network-based structural assurance* on trusting behaviors in online social networks. These constructs are refined based on related theories on face-to-face trusting behaviors, and tailored to the context of online social networks. Second, I introduce two new behavioral antecedents of trust—*trustor judgment bias* and *trustor status*—that have not been studied before, neither in face-to-face nor in online trusting behaviors. Third, I introduce novel operationalization methods to measure the behavioral trust-inducing factors in online social networks. Previous studies in online settings (e.g., e-commerce) that have used these constructs operationalized them with survey-based measures. The survey-based method has limited appeal in the context of online social networks due to the large number of members and the vast amount of data. This context requires an innovative method for measuring behavioral characteristics of members of online social networks. In this vein, I operationalize the behavioral trust-inducing factors using network measures, such as node (in/out) degrees, centrality measures, HITS authority, and PageRank. Fourth, I redefine and integrate existing graph theories and network-based concepts to develop a model of trust formation in online social networks, thereby enabling us to also study the interactions between behavioral trust-inducing factors and graph-based factors. Fifth, I differentiate between node importance and node prestige in signed social networks. Previous studies interchangeably used node prestige and node importance, and ignored their distinction in the context of signed networks. I argue that importance (general reputation or centrality) measures how often a node is connected to others (the number of relationships), while prestige considers the net value of incoming links to a node (the sign of the relationships).

This research also offers practical implications for social network management and applications. Identifying truly trusted members of a social network can benefit both network members and the community governance. Trusted members of the community are valuable in maintaining a sustainable functionality of the network. Trusted members of a social network are usually identified by other members of the community (e.g., by a voting process). Findings of this study can benefit in identifying truly trusted members of social networks.

In this study, I showed that not all the trusting votes are unbiased and the trustor's judgment bias should be considered in analyzing the trustworthiness of members. Rooted in psychological factors, such as personality, some users might be positively or negatively biased toward trusting others. For example, some people inherently can trust others, in general. Thus, in identifying trusted members of social networks, the bias of the voters (trustors) should be considered. Truly trusted members of the community can be identified by giving less weight to biased trustors' trust votes. This can be done by analyzing the user profiles and past relationship data (e.g., analyzing the ratio of blocked to trusted lists). Moreover, I found that reputable and central users are more likely to be perceived as being trustworthy and trusted by other members. Thus, the centrality of a user, the number of her interactions with others, and the extent to which the user is related to other important users should be considered in identifying trusted members.

All in all, the research findings indicate that the previous graph-based view of trust in social networks is not sufficient to understand trust formation among social network users.

The findings of this study can be used in identifying trustworthy reviews and aggregating crowd wisdom. Crowd wisdom can be aggregated by giving more weight to the opinions from trusted members. For example, in an online product review community, reviews written by trusted members should be given more weight than other reviews. Also, helpfulness or trustworthiness votes on the reviews from biased members should be given less weight in the aggregating of crowd wisdom. The findings of this study can also be used in friend recommendation systems by providing better recommendations based on users' status and similarity. I found that similar users who are in the same community with similar neighbors are more likely to trust one another or become friends. Moreover, users' status plays an important role in their relationships in social networks. I found that reputable (central) and high-status actors are more trusted by others. Therefore, both central (reputable) and high-status actors of the community can be recommended to new users for friendships or trust relationships.

The current study inevitably has some limitations, which may be addressed in future research. I could not control for users' demographic characteristics, such as age, gender, or even profile data, due to the absence of such information in the dataset. Also, I did not take into account users' previous interactions with one another, which might have had an effect on their trusting behaviors. The possible effects of users' demographic differences and previous interactions—when they are available—can be examined in future studies. I measured trusting beliefs from trustors' comments using human annotators, which is time-consuming. The model can be examined using a text mining system (to measure trusting beliefs) for a less costly and automated method. I tested the proposed conceptual model using one dataset from Wikipedia election for selecting trusted admins. The generalizability of the proposed conceptual model should be validated using trust data from other online social networks.

CHAPTER 3

Essay 2: Trust Prediction in Online Social Networks – A Theory-Based Predictive Model for Trust/Distrust Prediction in Online Social Networks

3.1. Introduction

The primary goal of social network analysis is to study the structure, relationships, dynamics, and interactions of online social networks, drawing on theories and methods from multiple disciplines, such as sociology, biology, computer science, mathematics, and statistics. Recent research has expanded insights into different aspects of social networks, such as information dissemination, social contagion and influence, link creation and network growth, and behavioral patterns in social networks.

Social networks are usually visualized as graphs, where a vertex represents a user (actor) and an edge represents some form of relationships between users. Forming new links among the users of social networks results in network growth. *Link prediction* in social networks is the task of determining if there will be a link or relationship between a pair of users, as well as the type of that relationship. Link prediction is one of the core problems in social network analysis (Backstrom and Leskovec 2011; Fang et al. 2013; Liben-Nowell and Kleinberg 2007; Wasserman and Faust 1994) and has recently received great attention. Prior research has mainly studied the link prediction problem in unsigned networks to predict future links between nodes of a social network. However, in many real social networks, e.g., Epinions, Slashdot, and Wikipedia, the sign (positive/negative, trust/distrust, friend/foe, etc.) of a connection is also an important part of the connection. Thus, in such signed networks, in addition to predicting link formation, the sign of a potential link has to be determined, giving rise to the *sign prediction* problem (Leskovec et al. 2010a).

Trust/distrust relationships are an important type of signed relationships in online social networks. Trust is a critical component of any interpersonal and social relationships in which risk, uncertainty, or interdependence exists (Fukuyama 1995; Gefen et al. 2003). In online interactions, trust plays a more critical role because of the inherent higher level of uncertainty, risk, and fear of opportunistic behaviors. Trust is a difficult decision in online communities because there are a large number of participants with different social backgrounds and perspectives (Ma and Agarwal 2007), and most users are anonymous or only limited information about them is publicly available. In spite of the critical role of trust in online relationships, trust in online social networks—how trust forms and what factors influence trusting behaviors—has not been well studied. Furthermore, most of the established theories, methods, and solutions for unsigned/undirected networks are not always applicable to the sign prediction problem in signed networks. For example, the network structural balance theory predicts balanced triads of undirected relationships, while relationships are directed in most of the real signed networks.

Previous research on trust in online social networks has mostly been in the field of computational network analytics. *Trust prediction* has been considered as a special case of sign prediction and existing knowledge and methods of link prediction have been used to build trust prediction models. Trust computational models use existing network structural data to predict future trust/distrust relationships. Similarity between users (Golbeck 2009; Kunegis et al. 2009; Matsuo and Yamamoto 2009; Tang et al. 2013; Ziegler and Golbeck 2007), and local and global network properties (Leskovec et al. 2010a; Leskovec et al. 2010b; Liu et al. 2008; Zhang et al. 2013; Ziegler and Lausen 2005) have been used as important factors in trust formation in social networks.

Some of the previous studies have tried to adapt general social graph theories to sign prediction using unsupervised models and techniques, such as the structural balance theory using clustering (Chiang et al. 2012), supervised learning (Chiang et al. 2011; Leskovec et al. 2010a), low-rank matrix factorization (Chiang et al. 2014; Tang et al. 2014), and collaborative filtering (Javari and Jalili 2014). However, previous studies have only used limited dimensions of trust formation among online users, resulting in an incomplete understanding of trust formation in online social interactions. Previous studies have used network structural dimensions (such as using only similarity or structural importance) to predict trust/distrust links. They have largely overlooked the underlying socio-psychological drivers in trust behaviors, such as the behavioral characteristics of the trustor and trustee. Furthermore, since distrust information usually is unavailable, previous trust prediction models predict only trust links (and not distrust links). This means that those models can predict only the existence or non-existence of trust links between users. Non-existence can mean either no relationship (neither trust nor distrust link) or a distrust link. In real relationships, there are distrust links and mixing them with no relationships is not appropriate.

I argue that network structural features should be used in trust/distrust prediction based on related behavioral theories of trust. Kane et al. (2014) argued that one of the challenges in studying online social networks or social media is behavioral. The main purpose of social media and online social networks is to support interpersonal communication and collaboration using

Internet-based platforms (Kane et al. 2014). It is obvious that users are still real people acting in online social media. Thus, the traditional antecedents of trust rooted in socio-psychological theories need to be considered in predicting trust/distrust links in online social interactions.

There are also characteristics and capabilities specific to an online social network that make it distinct from other settings. Online social networks are usually large, with more heterogeneity in the social characteristics of network members and more complexity in the structure (Wellman 1997). They differ from traditional social networks in terms of content and structure (Kane et al. 2014). Moreover, social media has provided its users novel ways of acting and interacting with each other that would have been difficult or impossible in earlier online or offline settings. In addition, centrality and connectedness of members of online social networks play a critical role in many aspects of sustainability of online social networks, such as information dissemination. These unique characteristics of online social networks may influence trust/distrust formation. This requires traditional antecedents of trusting behaviors to be adapted in context and used along with online social network-specific characteristics.

In this study, I strive to address the gaps in the existing literature by proposing a novel theory-based predictive model to predict trust and distrust links in online social networks. I propose a comprehensive set of predictors (both behavioral and structural) of online trust/distrust links, guided by relevant psychology, social science, and computational network theories. Based on these, I propose a supervised predictive model to predict trust/distrust links in social networks. To the best of my knowledge, this is the first theory-based trust prediction model of online social interactions. I empirically evaluate the utility of my proposed model using a real-world dataset. The empirical results show the superior fit and predictive performance of the proposed model over the baselines (previous models that used only limited structural dimensions of online social networks).

3.2. Related Work

3.2.1. Link Prediction

One of the interesting problems of social network analysis is to predict new link formation among people in the network (the link prediction problem). Link prediction, as a core problem in social network analysis, predicts the formation and type of future relationships among users based on available information. Social networks grow and change quickly over time, as new users and new relationships are added to the network. This dynamic nature and large size of social networks make the link prediction task a complicated problem.

Several heuristics and algorithms have been proposed to solve the link prediction problem. Previous methods can be widely categorized into unsupervised and supervised learning methods. Unsupervised methods compute scores for pairs of users based on some structural or topological properties of the network. Most previous unsupervised methods generate scores for similarity (proximity) of nodes, based on either common neighborhoods or paths between the two nodes. Then, future links are predicted based on the similarity scores. Some studies have used graph theories such as network structural balance theory to predict future links. For example, future links are predicted to make the network more balanced (Chiang et al. 2011; Leskovec et al. 2010a; Leskovec et al. 2010b).

Similarity-based methods, grounded in homophily (McPherson et al. 2001), are the most widely used unsupervised methods. The idea of homophily is that similar users are more likely to interact and connect with each other. Accordingly, higher similarity indicates a higher chance of linkage between a pair of users. Multiple local and global structural similarity measures have been proposed and used to measure the similarity between users. Liben-Nowell and Kleinberg (2007), as one of the first studies of link prediction, formalized the link prediction problem and developed several graph-based similarity measures for the “proximity” of nodes to predict future links. Local similarity measures are mostly based on common (mutual) neighbors between users (Liben-Nowell and Kleinberg 2007; Newman 2001), such as Jaccard coefficient and Adamic–Adar index. Global similarity measures, such as Katz index (Katz 1953), consider paths connecting users.

Supervised methods learn the underlying structure of link formation from observed links, and then predict the likelihood of future links by using the learned model. Supervised machine learning methods build a training data set from currently observed links, where each record includes a set of features (predictors) and a class variable (e.g., link or no-link). Multiple features, including network structural features, such as similarity and distance features (Al Hasan et al. 2006; Lichtenwalter et al. 2010; O’Madadhain et al. 2005), and demographical and geographical characteristics (O’Madadhain et al. 2005; Wang et al. 2011), have been used for link prediction. O’Madadhain et al. (2005) used content-based features, such as the divergence of the topic distributions of the two nodes, geographic proximity, and similarity of journal publication patterns, to predict the interaction between users. Al Hasan et al. (2006) used multiple features, including network topological features, aggregated features, and semantic similarity features, to predict link formation in a co-authorship network. Wang et al. (2007) used topological and content-based similarity measures to predict co-authorship relations. Wang et al. (2011) found that similarity between two individuals' movements strongly correlates with their proximity in the social network. Based on that, a supervised link prediction method was proposed using mobility and network-based similarity measures, such as common neighbors, co-location, and spatial cosine similarity. Hopcroft et al. (2011) proposed a factor-based graph model to predict reciprocal relationships on Twitter and found that elite users tend to follow each other, two-way relationships on Twitter are balanced, but one-way relationships are not. Backstrom and Leskovec (2011) proposed a supervised learning algorithm based on supervised random walks for link prediction, using network structural features, such as communication features, similarity, and common neighbors.

There are several disadvantages with unsupervised methods. Unsupervised models use predefined scores that are invariant to the specific structure of the input graph (Menon and Elkan 2011), do not involve any learning, and use only a single metric to predict link formation between users (Lichtenwalter et al. 2010). On the other hand, as supervised methods predict link formation based on a vector of features, they can also capture important interdependency relationships among the multiple features (Lichtenwalter et al. 2010). Compared to supervised link prediction methods, unsupervised link prediction methods are simpler and hardly applicable in sign prediction (where the focus is not only on network growth but also predicting the sign of links). In this study, I focus on supervised prediction methods to be able to predict both trust and distrust links.

3.2.2. Sign Prediction

In many real social networks, relationships carry a meaning, such as the sentiment of individuals toward each other, more than just the existence of the connection. Social networks in which relationships (edges) have positive or negative meaning are called signed social networks. Most previous studies focused on link prediction where the task is to predict the existence of a relationship, regardless of the type of the relationship. In a signed network, in addition to predicting the likelihood of a link, predicting the sign of the link is also of great importance. However, most of the theories and existing solutions for unsigned networks are not applicable to signed networks. For example, Chiang et al. (2012) showed that spectral clustering algorithms for unsigned networks cannot be directly used on signed networks.

Multiple theories and models, such as the theories of network structural balance and status (Chiang et al. 2011; Leskovec et al. 2010a; Leskovec et al. 2010b), have been applied to solve the sign prediction problem. Leskovec et al. (2010b) proposed a set of features based on 16 distinct signed directed triads of the balance theory in signed networks. Chiang et al. (2014) also interpreted Katz measures from the structural balance perspective for sign prediction. They proposed three sign prediction methods based on measures of social imbalance, higher-order cycles, and low-rank modeling. In other related work, Vu et al. (2013) addressed the sign prediction problem as a decision-making problem. They defined a decision making feature called Positive-Negative Ratio (PNR) based on local information of nodes. Javari and Jalili (2014) used a collaborative filtering approach to sign prediction problem and modeled a signed network as a bipartite user-item network.

Trust Prediction

Previous research has considered trust prediction as a specifial case of sign prediction. Trust is a matter of great importance in any type of online communication. In particular, trust has received remarkable attention in online social networks. Trust is a difficult decision in online social networks because most users are anonymous or only limited information is publically available. Some social networks offer an option to users to create a “web of trust” community from others whom they trust. They also can create a “block list” of others whom they distrust.

Computational trust prediction models use existing knowledge and methods of sign prediction to predict trust/distrust links. Information from users’ profiles and previous trusting behaviors, such as existing trust/distrust links (e.g., “web of trust” and “block list”), are used to predict future relationships. Network structural features and node properties, similarity, interactions, and contextual features have been used to build trust prediction models (Liu et al. 2008; Tang et al. 2013; Zhang et al. 2013). Multiple studies have found a strong correlation between users’ similarities (such as similarity of profiles) and interpersonal trust in social networks (e.g., Golbeck 2009; Ziegler and Golbeck 2007). Tang et al. (2013) studied the effect of homophily (similarity) on trust prediction and proposed an unsupervised framework by using low-rank matrix factorization for trust prediction. It has been empirically validated that similar users tend to trust one another and also trusted users are more similar. In another related stream of research on trust, trust propagation studies (e.g., Guha et al. 2004; Ziegler and Lausen 2005) showed the transitive characteristic of trust. Trust transitivity means if user i trusts user j, and user j trusts user k, then user i can be inferred to trust user k to some extent. For example, Guha et al. (2004) developed a framework of trust propagation schemes in different circumstances. In addition, distrust was incorporated in a computational trust propagation setting, and a formal and computational treatment of distrust propagation was developed. Distrust has been found to have significant effects on how trust propagates through the network.

Previous research has also used supervised learning to predict trust/distrust links. Liu et al. (2008) proposed a supervised classification method for trust prediction using a set of relevant features from user attributes and user interactions. They developed a trust taxonomy to be used in a binary classifier including two main categories, user factors (such as the number of reviews, the length of the review, or commenting frequency) and interaction factors (such as the number of a user's comments to another user's review). Matsuo and Yamamoto (2009) found that trust and opinion have strong mutual effects and observed the community gravity effect. A supervised prediction model was proposed based on the similarity of users' profiles, product ratings, and trust relations. Zolfaghar and Aghaie (2012) proposed a supervised trust prediction model using contextual and structural trust-inducing factors, such as similarity-based, reputation-based, and relationship-based factors. Zhang et al. (2013) proposed supervised machine learning algorithms to predict trust in social networks using a set of extended variants of commonly adopted local importance measures with combination of the idea of longer cycles and triangle patterns. Tang et al. (2014) studied trust/distrust propagation in social media by leveraging data mining and machine learning techniques.

Table 3.1- A Summary of Previous Studies and the Current Study

Study	Method	Theories	Independent Variables	Dependent Variables	Findings
Chiang et al. (2011)	Supervised machine learning	Network structural balance theory	Features derived from longer cycles in the network (measures of network social imbalance)	Link sign	The proposed supervised method outperforms all previous approaches
Chiang et al. (2012)	K-way clustering	Network structural balance theory	A criterion that is analogous to the normalized cut, called balance normalized cut	Link sign	Formulated new k-way objectives and kernels for signed networks and developed a multilevel clustering algorithm for signed networks
Chiang et al. (2014)	Unsupervised- Low rank modeling (matrix completion) approach	Network structural balance theory	General measures of social imbalance (MOIs) based on l-cycles in the network	Link sign	Global viewpoint of structural balance resulted in superior performance and computational gains in sign prediction and clustering
Golbeck (2009)	Unsupervised- a breadth-first search algorithm		Profile similarity features	Trust links	Showed that in addition to overall similarity, there is also correlation between trust and the largest single difference in ratings, and between trust and the agreement on movies the source has given extreme ratings. A composition of these measures predicts trust with higher accuracy, and less variation than when using overall agreement alone.
Guha et al. (2004)	Unsupervised propagation algorithms	Graph theory	Current trust and distrust relations (matrices)	Trust propagation	Developed a formal framework of trust propagation schemes, introducing a formal and computational treatment of distrust propagation

Javari and Jalili (2014)	Clustering and collaborative filtering algorithm	Network structural balance and graph theories	Balanced clusters and similarity between clusters	Link Sign	Proposed method outperformed previous methods
Kunegis et al. (2009)	Supervised machine learning	Graph theory	Various signed spectral similarity measures (signed Laplacian similarity matrix)	Link sign	Signed networks exhibit multiplicative transitivity that can be summarized by the phrase the enemy of my enemy is my friend.
Leskovec et al. (2010a)	Supervised machine learning (logistic regression)	Network structural balance and status theories	Local degree features, node status, and triad features (triangle relationships)	Link Sign	Model performance significantly improved over previous approaches; showed that employing information about negative relationships can be useful even for tasks that involve only positive relationships; Compared the learned models to theories of balance and status to find where they were consistent or inconsistent with these theories.
Leskovec et al. (2010b)	Probability model	Network structural balance and status theories	Triadic structures and node status	Link Sign	Investigated two theories of signed social networks (balance and status) in signed online networks; then developed an alternative theory of status that better explains the observed edge signs
Liu et al. (2008)	Supervised machine learning		Features derived from user attributes (review, rating, comment related) and user interactions (such as write-comment and write-rating or rate-rate connections)	Trust links	NB and SVM classifiers outperform the baseline classifier (that randomly assign 25% instances as positive)

Matsuo and Yamamoto (2009)	Supervised machine learning	Homophily	Similarity measures (profile, rating, neighbors)	Trust links	Community gravity (which is the bidirectional effect of trust and rating) was observed by investigating product propagation networks
Tang et al. (2013)	Unsupervised- low-rank matrix factorization with homophily regularization	Homophily	Similarity measures on user ratings	Trust links	Demonstrated the existence of homophily in trust relations and an unsupervised framework was proposed to capture its effect in trust relations
Tang et al. (2014)	Trust propagation and low-rank matrix factorization methods	Balance theory	Current trust and distrust relations	Trust links	Studied distrust in social media from the computational perspective and found that distrust is not the negation of trust and distrust has added value over trust.
Vu et al. (2013)	Supervised machine learning (logistic regression)	Decision making theory	Positive-negative ratio feature (PNR)	Link sign	Better classification accuracy and AUC than previous methods
Zhang et al. (2013)	Supervised machine learning		Local bias and PageRank	Link sign	Proposed a modified version of the PageRank algorithm for signet networks. An edge sign predictor using supervised machine learning algorithms was also established which significantly outperformed previous ones.
Ziegler and Golbeck (2007)	Probability model (correlation)		Profile similarity	Trust links	Showed the dependencies between trust and user similarity exists; and as trust between users increases, the difference in the ratings they assign to movies decreases

Ziegler and Lausen (2005)	Unsupervised- an algorithm based on spreading activation models		Local group trust metrics	Trust and distrust propagation	Introduced a classification scheme for trust metrics along various axes and discussed advantages and drawbacks of existing approaches for Semantic Web scenarios; proposed Appleseed for local group trust computation.
This study	Supervised machine learning	Socio-psychological theories of trust (personality-based, institution-based, cognition-based trust); Social status theories (expectation states theory, status characteristics theory); Homophily; Graph theory	Trustee features (prestige and importance); Trustor features (status and trust propensity); Similarity features; Status differential features	Trust and distrust links	Propose a theory-based predictive model of trust and distrust links in online social network The proposed models show superior performance over the baselines (previously used models)

3.3. Research Questions

Most previous studies on trust/distrust prediction in online social networks are in the field of computational network studies. Table 3.1 presents a summary of related previous studies. Most of them focused on building computational models using very limited network structural information and graph concepts. Most of them considered only one factor, e.g., similarity measures, ignoring any other underlying motives in social relationships. In addition, previous models did not consider any underlying behavioral theories of trust formation among users. Some studies have used social graph theories such as structural balance, which limited the link prediction to one aspect of relationships.

Structural balance focuses on triadic relations and posits that my friend’s friend is my friend, my friend’s enemy is my enemy, my enemy’s friend is my enemy, and my enemy’s enemy is my friend. Prediction models based on structural balance try to predict links to make the network more balanced, ignoring any other motives or related factors. Furthermore, since distrust information is usually not available, previous models are limited to predict trust or no-trust links (not distrust links distinctly). That also means that the valuable information from distrust relationships was ignored in predicting trust links.

As discussed earlier, distrust has been found to be very valuable in predicting trust propagation. Factors inducing distrust are more decisive than trust since people are more certain when distrusting others. Thus, ignoring distrust in trust studies leads to an incomplete and biased estimate of trust links. Furthermore, evidence from socio-psychological studies suggests that trust has behavioral antecedents, such as individual’s cognitive process, inherent propensity to trust (Mayer et al. 1995; McKnight et al. 1998), situational factors such as structural assurance

(McKnight et al. 1998), and social status of individuals (Lount Jr. and Pettit 2012; Ridgeway and Walker 1995). The behavioral trust-inducing factors were not examined in the previous trust prediction models. To address the discussed gaps in the existing literature, I specifically try to answer the following research questions in this study:

- What are the important predictors of trust/distrust links in signed online social networks based on related traditional trust theories (theory-based behavioral predictors)? And, how can they be adapted and measured in the context of online social networks?

- What are the important predictors of trust/distrust links in signed social networks based on the specific characteristics of online social networks (graph-based structural predictors)?

- How can the identified predictors be combined to build a supervised trust prediction model in signed online social networks? And, how do they together contribute to the performance of the predictive model?

To answer the research questions, I propose a theory-based trust/distrust prediction model based on related theories of social psychology and social graphs, and by using both trust and distrust information. I use structural and behavioral predictors of trusting behaviors, including trustor and trustee characteristics, situational characteristics, status related characteristics, and similarity. Based on the proposed model, I build multiple supervised classification models to predict trust/distrust links in online social networks.

3.4. Proposed Predictive Model

In this section, the theoretical foundations of the proposed model are explained. Then, based on the theoretical trust predictors, I build the predictive model of trust/distrust prediction in social networks. As discussed in Essay1, trust is a multi-dimensional concept and has been studied in multiple contexts. Trust is defined as a trustor's willingness to be vulnerable to the actions of a trustee based on positive expectations of the trustee's intentions (Mayer et al. 1995; Mcknight et al. 1998). Trust has been conceptualized according to several theoretical streams, including knowledge-based trust, calculative-based trust, institution-based trust, cognition-based trust, and personality-based trust (Gefen et al. 2003; Mayer et al. 1995; Mcknight et al. 1998).

Based on these behavioral theories and the related concepts and measures of social graphs, I define a comprehensive set of trust-inducing features to use in supervised machine learning algorithms. I consider the trust prediction problem as a binary classification problem, where a potential link is classified as either a trust link or a distrust link. A trust link forms when a trustor trusts a trustee, and a distrust link forms when a trustor distrusts a trustee. For example, in some online review websites, users can build a web of trust composed of reviewers whose reviews they trust or block reviewers whose reviews they distrust. When user *a* adds user *b* to her web of trust, it means user *a* (the trustor) issues a trust link toward user *b* (the trustee). When user *a* adds user *b* to her block list, it means user *a* (the trustor) issues a distrust link toward user *b* (trustee).

3.4.1. Trust/Distrust link Predictors

Trustor Network-Based Trust Propensity

As discussed in essay 1, trustor's *trust propensity* (also called *dispositional trust*) is one of the important antecedents of trusting behaviors and is an individual's inherent characteristic (Mayer et al. 1995; McKnight et al. 1998). Trustor's trust propensity has been found to be one of the important factors in initial trust formation where no previous experience exists among parties (Mayer et al. 1995; McKnight et al. 1998; McKnight and Chervany 2001; Williams 2001). In the context of online relationships that are distant and where limited information is available about the parties, disposition to trust could play an important role in initial trust formation (McKnight et al. 2004). This effect can be even more important in online social networks, where there is even less information about users and no previous experiences usually exist.

The trustor's trust propensity is an inherent characteristic of the trustor to trust others irrespective of any specific trustee. The more a trustor tends to trust others easily, the more trust links (rather than distrust links) will be issued toward others. Therefore, a trustor's propensity to trust can be measured by the proportion of the trust links to the total number of links issued toward others.

I define a feature set for trustor propensity predictor using the trustor's issued links toward others. First, I consider the proportion of the total trust links to the total issued links toward others: $PropensityP(i) = \frac{outdegree^{+}(i)}{outdegree^{+}(i)+outdegree^{-}(i)}$, where $outdegree^{+}(i)$ is the out-degree of positive (trust) links of user i (the trustor) in the network, and $outdegree^{-}(i)$ is the out-degree of negative (distrust) links of the user. Second, I take into account both trust and distrust links: $PropensityF(i) = \frac{outdegree^{+}(i)-outdegree^{-}(i)}{outdegree^{+}(i)+outdegree^{-}(i)}$. Third, I use a sigmoid function of trust links (as used in Zolfaghar and Aghaie, 2012): $PropensityZ(i) = \frac{1}{1+ e^{-\propto(|outdegree^{+}(i)|-\mu)}}$; a sigmoid function is used to keep the value in the range [0, 1]; α is set to 0.1 to control the slope and μ is set to 0 to control the midpoint of the sigmoid curve.

Trustee Network-Based Importance (Structural Assurance)

As discussed in essay 1, *institution-based trust* (also called *impersonal trust*) has been found to be an important antecedent of trusting behaviors and has two types of situational normality and structural assurance (McKnight et al. 1998). In the context of this study, I assume that situational normality is the same for everyone since all users are acting in the same online community. Structural assurance increases trust and affects trusting beliefs by giving the sense that parties in the situation are trustworthy (Gefen et al. 2003; Mcknight et al. 1998). In essay 1, I argued that a user's position in the network could be an indicator of the user's reputation. And, network structural reputation measures (such as *centrality* measures (Katz 1953), *PageRank* (Page et al. 1999) and its variants, and *HITS authority* (Kleinberg et al. 1999)) have been used to predict influential and trusted users in online social networks (Song et al. 2007; Varlamis et al. 2010).

In this study, I consider the public view of a trustee as the general assurance of the trustee's trustworthiness and as a predictor of forming a trust/distrust link with other users. I define a feature set for structural assurance predictor (or the trustee's general reputation in the network) using centrality measures, HITS authority, PageRank, and clustering coefficient by incorporating all the links (including trust and distrust links) issued by all actors. By including all the links, I could examine how central and important the trustee is in the network. Centrality measures (Freeman 1978) identify key nodes (users) for information dissemination in a social network, by providing leadership or bridging different communities (Chau and Xu 2012). *Closeness centrality* means nodes at the geographic center are central and measures how long it takes for information to pass from a trustee node to other nodes in the network. It is based on the length of the average shortest path between a node and all nodes in a social graph (Bavelas 1950): $ClosCent(a) = \frac{N-1}{\sum_{b \in N} d(a,b)}$, where *d(a,b)* is the distance between nodes *a* and *b*, and *N* is the number of nodes in the graph. *Betweenness centrality* means nodes with many transits are central and measures how often a trustee node is found on the shortest path between two other nodes in the network. Trustee betweenness centrality measures how often a trustee node *a* is found on the shortest path between two other nodes in the network (Brandes 2001): $BetwCent(a) = \sum_{i,j \in N, i \neq j \neq a} (\frac{\sigma_{i,j}(a)}{\sigma_{i,j}})$, where $\sigma_{i,j}(a)$ is the number of shortest paths from *i* to *j* that pass through node *a*, and $\sigma_{i,j}$ is the total number of shortest paths from node *i* to node *j*. Hyperlink-Induced Topic Search (HITS) is a link analysis algorithm for rating web pages and consists of two measures: *hub* and *authority* (Kleinberg et al. 1999). A good hub represents a page that points to many other pages, and a good authority represents a page that is linked by many different hubs. In other words, a hub is a node with many out-links and an authority is a node with many in-links. Authority of a node *a* and hub of a node *a* are iteratively calculated as follows (Kleinberg 1999): $HitsAuth(a) = \sum_{a,b \in E} HitsHub(b)$, $HitsHub(a) = \sum_{a,b \in E} HitsAuth(b)$; where *E* is the set of directed edges (links) in the social graph.

Here, I only consider trustee authority, since a well-known trustee is one with many in-links rather than out-links. In the context of social graphs, PageRank is an iterative algorithm that measures the importance or authority of each node within the network. It has been discussed that a user (node) is influential if other influential users trust or follow her in a trust social network

(Varlamis et al. 2010). The PageRank score of user a is iteratively calculated as follows (Page et al. 1999): $PageRank(a) = \frac{(1-d)}{N} + d\sum_{b \in |indegree_a|} \frac{PageRank(b)}{|outdegree_b|}$; where d is the so-called damping factor, and *N* is the number of nodes in the graph. Clustering coefficient measures the localized density or how close the neighborhood (the set of nodes immediately adjacent to node a) of a node a is to a complete subgraph. That means nodes which are located in a dense neighborhood of the network are more likely to grow than those in a sparse neighborhood (Al Hasan et al. 2006). For a directed network, it is calculated as follows (Watts and Strogatz 1998): $ClusCoef(a) = \frac{\left|\left\{(i,j) \in E \middle| (a,i) \in E \wedge (a,j) \in E\right\}\right|}{K_a \cdot (K_a - 1)}$; where E is the set of edges (links), and k_a is the total degree (in&out-degree) of node a.

Network-Based Social Status

Social *status* (Berger et al. 1998) of both trustee and trustor have been found to influence trusting behaviors.

Trustee Network-Based Status

As discussed in essay 1, high-status actors of the group are believed to be more worthy, competent, and influential (Ridgeway and Walker 1995). Higher status is also associated with greater perceived trustworthiness and higher levels of trust (Cook et al. 2009).

In the context of online social networks, social status is defined as the position or rank of a user in the network and represents the degree of honor or prestige attached to that position. In other words, status is the aggregate opinion of others about a specific node in the network. A node's status (prestige) in a network has been considered as an important predictor of received positive links (friendship or trust) from other nodes of the network (Kunegis et al. 2009). A more prestigious user is one who is more trusted by others. A user's prestige (status) in online social networks is often identified by incorporating both trust and distrust links toward the user, such as the *Fans Minus Freaks* (FMF) (Kunegis et al. 2009) measure, where fans are ones who trust the user and freaks are ones who distrust the user. Trusted users have a high number of fans and distrusted users a high number of freaks. Therefore, a trustee with a higher status is perceived to be more trustworthy and receives more trust links from other members of the network.

I define a feature set for trustee status predictor (or the trustee's prestige in the network) using the issued links from others to the trustee. I consider FMF (*StatusFM*), the ratio of FMF to the total indegree: $StatusFS(i) = \frac{indegree^{+}(i) - indegree^{-(i)}}{indegree^{+(i)} + indegree^{-}(i)}$, and to the total degree of the trustee: $StatusT(i) = \frac{indegree^{+}(i) - indegree^{-(i)}}{degree\ (i)}$. I define two measures based on the incoming trust links as: $StatusPR(i) = \frac{indegree^{+}(i)}{indegree^{+(i)} + indegree^{-}(i)}$ and $StatusPO(i) = \frac{indegree^{+}(i)}{degree\ (i)}$. I also use the prestige measures by Zolfaghar and Aghaie (2012) using trust links: $StatusZ(i) = \frac{1}{1 + e^{-\propto(|indegree^{+}| - \mu)}}$, where a sigmoid function is used to keep the value in the range [0, 1]; α is set to 0.1 to control the slope and μ is set to 0 to control the midpoint of the sigmoid curve. Leskovec's measure of status is also used (Leskovec et al. 2010a; Leskovec et al. 2010b), in which a trustee node's status increases for each positive link it receives and each negative link it issues, and declines for each negative link it receives and each positive link it issues: $StatusL(i) = \frac{indegree^{+}(i) - indegree^{-(i)} - outdegree^{+}(i) + outdegree^{-(i)}}{degree(i)}$. Furthermore, *Net Trust Votes (NTV)* is a measure of the Shapley value based centrality measures (Aadithya et al. 2010; Gangal et al. 2016) for directed signed networks. Since the main idea of the NTV is the same as the (Leskovec et al. 2010a; Leskovec et al. 2010b) status measure, I also considered it as a measure of trustee status: $StatusN(i) = \frac{1}{2}\left(indegree^{+}(i) - indegree^{-}(i)\right) - \frac{1}{2}\left(outdegree^{+}(i) - outdegree^{-}(i)\right)$.

Trustor Network-Based Status

It has been found that status also influences trusting others, such that individuals with higher status show more initial trust than individuals with lower status (Lount and Pettit 2012). Individuals with higher status have expectations that others will have favorable motives and display positive behaviors toward them, such as respect and praise. High social status brings the individual a set of internalized beliefs and expectations around the rewards she will receive by virtue of this elevated social position (Berger et al. 1998). The reward expectations involve social rewards, which induce positive expectations about others' motives and behaviors. Thus, trustors with higher status judge others' intentions as more benevolent, and in turn, trust others more easily (Berger et al. 1998; Lount Jr. and Pettit 2012). I posit that this could also be true for online social interactions. I use the same status features defined for trustee status as the feature set for trustor status predictor.

Network-Based Status Differential

I argued in essay 1 that one of the important concepts related to status is its "relative" characteristic. And, as status is relative in groups, perceived trustworthiness of individuals is also a relationship-level attribute. The *status differential* between two actors will determine the degree to which an actor perceives trustworthiness qualities in the other actor (Campos-Castillo and Ewoodzie 2014). Therefore, the status differential between a trustee and a trustor has an important role in shaping the trustor's beliefs about the trustee's trustworthiness. As the trustor's status relative to the trustee decreases (the status differential increases), the trustor's perceived trustworthiness of the trustee increases. Moreover, the *Status Theory*, as one of the online social network theories, was developed for directed social networks to better explain the observed edge signs and provides insights into the underlying social mechanisms (Leskovec et al. 2010a; Leskovec et al. 2010b). Based on the Status Theory, a trust (distrust) link from a trustor to a trustee indicates that the trustee has a higher (lower) status than the trustor (Leskovec et al. 2010a; Leskovec et al. 2010b). This also reinforces the discussed relational characteristic of status by previous socio-psychological scholars.

Status differential between a trustee and a trustor is calculated as the difference between the trustee's status and the trustor's status: *StatusDifferential* (*trustee, trustor*) = *Status*(*trustee*) – *Status*(*trustor*). I use status features, $StatusFS$, $StatusT$, $StatusPR$, and $StatusL$, to obtain the status differential features.

Network-Based Homophily (Similarity)

As discussed in essay 1, *homophily* has been found as an important factor in social network formation, such as in friendship networks (McPherson et al. 2001). Homophily has been also observed and studied in online social network relationships (e.g., Wang et al. 2011; Xiang et al. 2010). Furthermore, homophily (i.e., similarity among users) has been used in link formation and sign prediction (predicting trust/distrust or friend/foe relationships) in online social networks. Multiple studies have shown that similarity is an important predictor of trust formation among users in online social networks (Golbeck 2009; Kunegis et al. 2009; Symeonidis et al. 2010; Tang et al. 2013; Zheng et al. 2014). It has been found that having a shared group membership enhances trustworthiness perceptions, thereby increasing the levels of initial trust (Meyerson et al. 1996; Tanis and Postmes 2005).

In the context of social networks, a shared group membership can translate into being in the same community with common neighbors. Users of the same group or community tend to perceive themselves as more trustworthy than others not in their group, and trust one another (Mcknight et al. 1998). Common links (neighbors) have also been widely used to measure the similarity between users in social networks, such as in the *Adamic-Adar index*, and in the *Jaccard coefficient* (Adamic and Adar 2003; Golder and Yardi 2010; Liben-Nowell and Kleinberg 2007; Lichtenwalter et al. 2010). Accordingly, I consider common neighborhood as a similarity measure among users. Thus, a trustor's trust in a trustee increases as the trustor has more common neighbors with the trustee.

I use three commonly used similarity measures in social networks, Jaccard similarity, Dice similarity, and inverse log-weighted similarity (Adamic and Adar 2003), to define the feature set for similarity predictor. The Jaccard similarity coefficient is the number of matches of two sets divided by their union or the number of common neighbors divided by the number of nodes that are neighbors of at least one of the two nodes. The *Dice similarity* coefficient of two nodes is twice the number of common neighbors divided by the sum of the degrees of the nodes. The *inverse log-weighted similarity* of two nodes is the number of their common neighbors, weighted by the inverse logarithm of their degrees.

3.4.2 Model

The proposed predictive model is shown in Figure 3.1, and the proposed predictors with their feature sets are summarized in table 3.2.

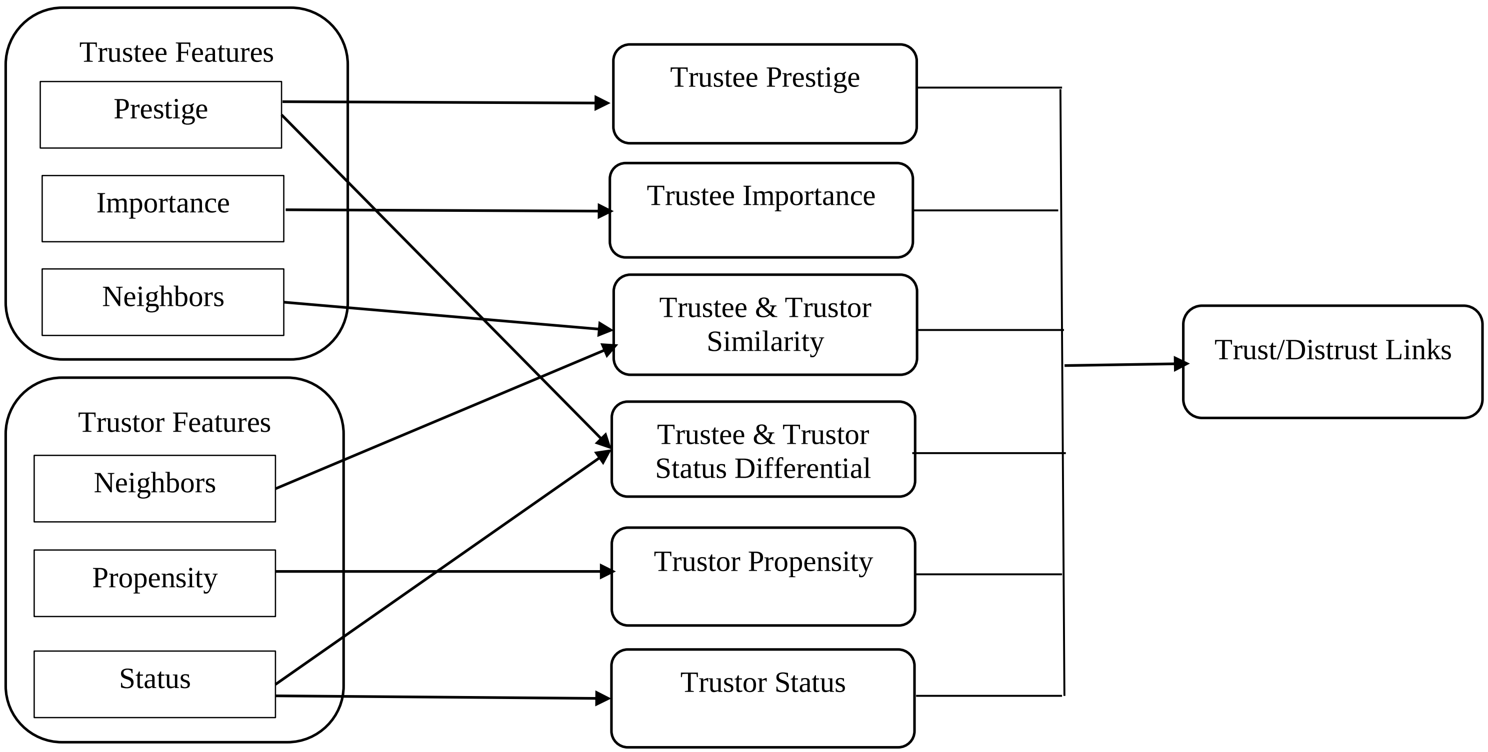

Figure 3.1- The Proposed Model

Table 3.2- The Proposed Feature Sets

Type		Theoretical Construct	Related References	Feature	Description
Trustee Features	Trustee Prestige	Social Status	Berger et al. 1998; Cook et al. 2009; Kunegis et al. 2009; Ridgeway and Walker 1995	StatusFM	The difference between positive indegree and negative indegree
				StatusFS	The difference between positive indegree and negative indegree divided by the total indegree
				StatusT	The difference between positive indegree and negative indegree divided by the total degree
				StatusPR	The positive indegree divided by the total indegree
				StatusPO	The positive indegree divided by the total degree
				StatusZ	The sigmoid function of the total positive indegree
				StatusL	The sum of the positive indegree and negative outdegree minus the sum of the negative indegree and positive outdegree, divided by the total degree
				StatusN	The Net Trust Votes (NTV) based on the Shapley value centrality
	Trustee Importance	Structural Assurance	Gefen et al. 2003; Katz 1953; Kleinberg 1999; McKnight et al. 1998; Page et al. 1999;	ClosCent	How long it takes for information to pass from the trustee to other nodes
				BetwCent	How often the trustee is found on the shortest path between two other nodes
				HitsAuth	How much the trustee is a good authority or is linked by many different nodes
				PageRank	The importance of a trustee within the network based on its influential followers (trustors)
				ClusCoef	The localized density or how close the neighborhood of the trustee is to a complete sub-graph

Trustor Features	Trustor Status	Social Status	Berger et al. 1998; Lount Jr. and Pettit 2012; Ridgeway and Walker 1995	StatusFM	The difference between positive indegree and negative indegree
				StatusFS	The difference between positive indegree and negative indegree divided by the total indegree
				StatusT	The difference between positive indegree and negative indegree divided by the total degree
				StatusPR	The positive indegree divided by the total indegree
				StatusPO	The positive indegree divided by the total degree
				StatusZ	The sigmoid function of the total positive indegree
				StatusL	The sum of the positive indegree and negative outdegree minus the sum of the negative indegree and positive outdegree, divided by the total degree
				StatusN	The Net Trust Votes (NTV) based on the Shapley value centrality
	Trustor Trust Propensity	Trust Propensity	Mayer et al. 1995; McKnight et al. 1998; Mcknight et al. 2004; Williams 2001	PropensityP	The positive outdegree divided by the total degree
				PropensityF	The difference between positive outdegree and negative outdegree divided by the total outdegree
				PropensityZ	The sigmoid function of the total positive outdegree
Status Differential Features		Social Status	Berger et al. 1998; Campos-Castillo and Ewoodzie 2014; Leskovec et al. 2010a; Leskovec et al. 2010b; Ridgeway and Walker 1995	StatusFS_Diff	The difference between the StatusFS values of the trustee and the trustor
				StatusT_Diff	The difference between the StatusT values of the trustee and the trustor
				StatusPR_Diff	The difference between the StatusPR values of the trustee and the trustor
				StatusL_Diff	The difference between the StatusL values of the trustee and the trustor

Similarity Features	Homophily	Adamic and Adar 2003; Liben-Nowell and Kleinberg 2007; McPherson et al. 2001	Jaccard_Sim	The number of common neighbors divided by the number of their neighbors combined together
			Dice_Sim	Twice the number of common neighbors divided by the sum of the degrees of the nodes
			Inverse log-weighted_Sim	The number of their common neighbors, weighted by the inverse logarithm of their degrees

3.5. Experiment

In the following section, I describe the experiment design and evaluation criteria, and the real-world dataset used to evaluate the proposed predictive model. I then report the results.

3.5.1. Dataset

I examined the predictive power of the proposed trust/distrust prediction model using a real world dataset from Epinions.com, one of the largest online product review communities. This product review community provides the opportunity for its users to make a web of trust of those whom they trust and a block list of those whom they distrust. The dataset (Leskovec et al. 2010b, also available by Stanford Network Analysis Project (SNAP)) includes a total of 841,372 links, in which 717,667 are trust links and 123,705 are distrust links, among 131,828 users. I sampled 10,000 links from 1825 unique nodes to create a balanced sample of 5000 trust and 5000 distrust links.

3.5.2. Experiment Design and Evaluation

To answer the research questions, I address the trust prediction problem as a supervised classification problem. Based on the predictors (features) in the proposed model, I build the predictive model. Then, to compare the predictive power of the proposed model with that of previous models, I build a set of baselines using the feature types used in previous studies. The predictive power of the proposed model, as compared to the baselines, is examined using multiple supervised machine learning algorithms. Specifically, five categories of feature sets were built and examined, including similarity, status differential, trustee PageRank importance, trustee centrality, and trustee FMF prestige features.

It is very common to assess the predictive performance of a classification model using criteria such as model accuracy, precision, recall, and F-measure. It is also important to assess the descriptive ability of a classification model in addition to its predictive performance. Ritschard and Zighed (2003) argue that the descriptive classification model and the classifier itself should be distinguished, and the fit of descriptive classification model should be assessed in addition to the classifier predictive performance. The descriptive ability of a statistical model is referred to its *goodness of fit*, which measures how well the model fits a set of observations. Hence, I examined both goodness of fit and predictive performance of the proposed predictive model, in comparison to baseline models.

Goodness of Fit

In predictive models, the goodness of fit is usually measured by the discrepancy between observed and predicted values. Deviance is a commonly used measure to test the goodness-of-fit of a statistical model. The general idea of deviance of a statistical model is to measure how far the model is from the observed values. Deviance is measured by minus twice the log-likelihood of model M (-2LogLik(M)) (Agresti 2013). The Log-Likelihood Ratio statistic (LRT) has an approximate Chi-square distribution (Bishop et al. 1977). Deviance permits to test the difference between a model M1 and a restricted version M2 with the difference D(M2|M1) = D(M2) – D(M1), which has a Chi-square distribution with degrees of freedom equal to $df2 - df1$, where $df1$and $df2$ are the degrees of freedom for model *M1* and *M2*, respectively.

In this study, I defined multiple features for each predictor. To select the best parsimonious feature sets, I ran Log-Likelihood Ratio goodness of fit test. I tested if losing any of the features affects the goodness of fit. I compared the goodness of fit of the full model (which includes all the features) to other reduced (nested) models that include fewer features. If there is no significant difference in the model fit, the reduced model is selected. To compare two nested models, the model with the lower deviance is deemed to be better. I used the Log-Likelihood Ratio statistic, which is the difference of two nested models *M1* and *M2* with an associated p-value to compare the models in terms of fit. The results of LR goodness of fit test for the full model and the best parsimonious model are summarized in table 3.3. The fit (deviance) of the full model is not significantly better than the reduced model. Therefore, I selected the reduced model. The final feature sets based on the best parsimonious model is summarized in table 3.4.

Table 3.3- Comparison of the Full and Reduced Model

Model	**df**	**Deviance**	**LRT Statistic ($\chi 2$)**	**dfF-dfR**	**Sig.**
Full	32	3786.4			
Reduced	21	3802.04	15.64	11	0.155

Table 3.4- The Final Feature Sets

Predictor	***Features***
Trustee Prestige	StatusFS; StatusT; StatusPR; StatusPO
Trustee Importance	Betweenness_cent; PageRank; Clustering_coeff
Trustor Status	StatusFM; StatusFS; StatusT; StatusPO; StausZ; StatusL; StatusN
Trustor Propensity	PropensityP; PropensityF; PropensityZ
Status Differential	StatusPR_Diff
Similarity	Dice_Sim; Inverse log-weighted_Sim

Then, I compared the goodness of fit of the best parsimonious model to the baseline models. As demonstrated in table 3.5, the fit of the proposed model is significantly better than other previously used bassline models.

Table 3.5- Comparison of Goodness of Fit of the Proposed Model with Baselines

Model	Deviance	df	Test	LRT Statistic	dfi-dfj	Sig.
Proposed Model (P)	3802.04	21				
Similarity (Sim)	13860.79	2	P vs Sim	10074.39	30	<.001
Status Differential (StatDiff)	9241.77	2	P vs StatDiff	5455.37	30	<.001
Trustee PageRank Importance (PR)	12499.58	2	P vs PR	8713.18	30	<.001
Trustee Centrality (Cent)	13633.56	3	P vs Cent	9847.16	29	<.001
Trustee FMF Prestige (FMF)	13133.77	2	P vs FMF	13133.77	30	<.001
Similarity & Trustee Prestige	13082.7	3	P vs Sim&FMF	9296.3	29	<.001
Similarity & Trustee Importance & Centrality	12444.55	5	P vs Sim&PR&Cent	8658.15	27	<.001
Similarity & Trustee	11858.94	6	P vs Sim&FMF&PR&Cent	8072.54	26	<.001
Similarity & Trustee & Status Difference	8103.2	7	P vs Sim&FMF&PR&Cent &StatDiff	4316.81	25	<.001

Prediction Performance

I then examined the prediction performance of the proposed model. I designed six experiments to examine the predictive power of the proposed model compared to previous baseline models. I examined previously used predictors in different experiments. Specifically, I used similarity, trustee centrality, trustee PageRank importance, trustee FMF status, and status differential features. Furthermore, I built multiple predictive models by combining previously used predictors. Then, I used the feature set built on the proposed model including similarity, status differential, trustee importance, trustee prestige, trustor status, and trustor trust propensity features in the last experiment.

To test the generalizability of the effects of the feature type over different classification methods, I used five standard classification methods: Naïve Bayes (NB), Logistic regression (LR), Support Vector Machines (SVM), Neural Network (NN), and J4.8 Decision Tree (DT). Therefore, the experiment followed a 10 (feature type) * 5 (classification method) full factorial design. I used Weka (Hall et al. 2009), an open source data mining platform including a collection of machine learning algorithms. Standard performance measures (Accuracy, F-Measure, and MCC) were used to evaluate the performance of the classifiers. For each combination of feature type and classification method, I ran 10-fold cross validation 10 times, obtaining 100 estimates of each performance measure. Therefore, I got 10*5*100 = 5000 observations. Based on the classification confusion matrix, the performance measures are defined as follows:

Confusion Matrix		$\hat{\boldsymbol{y}}$ ***(Predicted)***	
		Trust	Distrust
$\boldsymbol{y}(\boldsymbol{Actual})$	Trust	TP	FN
	Distrust	FP	TN

$$\text{Accuracy} = \frac{TP + TN}{P + N}$$

$$\text{Precision} = \frac{TP}{TP+FP}; \text{Recall} = \frac{TP}{TP+FN}; \qquad \text{F} - \text{Measure} = \frac{2*Precision*Recall}{Precision+Recall}$$

$$\text{MCC} = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

3.6. Results

The results of the six experiments are summarized in Table 3.6 and contrasted in Figure 3.2. The proposed model achieved the best performance in terms of all three performance measures across all the five classification methods. The best performance in terms of all three measures was achieved by using the proposed model features with J4.8 decisions tree (0.95 accuracy, 0.95 F-measure, and 0.90 MCC).

Overall, baselines showed poorer predictive performance compared to the proposed model. The best performance of baseline models was achieved by using status differential features (0.78 accuracy, 0.77 F-measure, and 0.55 MCC). The worst accuracy (0.49)—essentially no difference from random guess—was obtained by using only similarity features.

To show the superiority of the proposed model and the importance of using all the proposed features, multiple models were created by combining previously used features. Among them, the best performance was achieved by combining similarity and trustee (prestige and centrality and importance) and status differential features and by decision tree classifier (0.82 accuracy, 0.82 F-measure, and 0.63 MCC). However, the performance was significantly lower than the proposed model performance.

Table 3.6- Results: Average Accuracy, F-Measure, and MCC

Feature Type	Classifier	Performance					
		Accuracy		F-measure		MCC	
		AVE	STDV	AVE	STDV	AVE	STDV
Similarity	Naïve Bayes	0.49	0.01	0.64	0.01	-0.04	0.03
	Logistic Regression	0.55	0.01	0.49	0.02	0.10	0.03
	Support Vector Machine	0.51	0.02	0.20	0.12	0.04	0.04
	Neural Network	0.50	0.00	0.35	0.34	0.03	0.04
	Decision Tree	0.58	0.01	0.66	0.04	0.19	0.04
Trustee Centrality	Naïve Bayes	0.54	0.01	0.24	0.03	0.13	0.03
	Logistic Regression	0.58	0.02	0.57	0.05	0.17	0.03
	Support Vector Machine	0.54	0.01	0.22	0.03	0.13	0.03
	Neural Network	0.58	0.02	0.52	0.13	0.17	0.04
	Decision Tree	0.62	0.01	0.65	0.02	0.24	0.03
Trustee Importance	Naïve Bayes	0.61	0.01	0.41	0.03	0.29	0.03
	Logistic Regression	0.64	0.01	0.56	0.02	0.30	0.03
	Support Vector Machine	0.63	0.01	0.49	0.02	0.31	0.03
	Neural Network	0.63	0.02	0.56	0.05	0.29	0.04
	Decision Tree	0.65	0.01	0.58	0.02	0.33	0.03
Trustee Prestige	Naïve Bayes	0.60	0.02	0.55	0.02	0.20	0.03
	Logistic Regression	0.59	0.01	0.62	0.02	0.18	0.03
	Support Vector Machine	0.59	0.02	0.62	0.01	0.17	0.03
	Neural Network	0.62	0.04	0.50	0.06	0.29	0.10
	Decision Tree	0.65	0.01	0.52	0.03	0.37	0.03
Status Differential	Naïve Bayes	0.77	0.01	0.77	0.01	0.55	0.02
	Logistic Regression	0.78	0.01	0.77	0.01	0.55	0.02
	Support Vector Machine	0.78	0.01	0.77	0.01	0.55	0.02
	Neural Network	0.77	0.01	0.76	0.02	0.54	0.03
	Decision Tree	0.78	0.01	0.77	0.01	0.55	0.03
Similarity & Trustee Centrality & Importance	Naïve Bayes	0.61	0.01	0.43	0.03	0.29	0.03
	Logistic Regression	0.64	0.01	0.55	0.02	0.31	0.03
	Support Vector Machine	0.63	0.01	0.49	0.02	0.31	0.03
	Neural Network	0.65	0.02	0.60	0.05	0.31	0.04
	Decision Tree	0.67	0.01	0.66	0.03	0.35	0.03
Similarity & Trustee Prestige	Naïve Bayes	0.55	0.02	0.62	0.01	0.11	0.03
	Logistic Regression	0.61	0.01	0.63	0.02	0.22	0.03
	Support Vector Machine	0.63	0.01	0.65	0.01	0.26	0.03
	Neural Network	0.68	0.02	0.63	0.04	0.37	0.04
	Decision Tree	0.71	0.01	0.69	0.01	0.42	0.03

Similarity & Trustee (Prestige & Centrality & Importance)	Naïve Bayes	0.61	0.01	0.42	0.03	0.29	0.03
	Logistic Regression	0.68	0.01	0.62	0.02	0.38	0.03
	Support Vector Machine	0.64	0.01	0.52	0.02	0.32	0.03
	Neural Network	0.70	0.02	0.67	0.04	0.41	0.03
	Decision Tree	0.73	0.01	0.73	0.02	0.47	0.03
Similarity & Trustee (Prestige & Centrality & Importance) & Status Differential	Naïve Bayes	0.73	0.01	0.66	0.02	0.50	0.02
	Logistic Regression	0.81	0.01	0.81	0.01	0.62	0.02
	Support Vector Machine	0.81	0.01	0.81	0.01	0.62	0.02
	Neural Network	0.81	0.01	0.82	0.01	0.63	0.02
	Decision Tree	0.82	0.01	0.82	0.01	0.63	0.02
The Proposed Model	Naïve Bayes	0.88	0.01	0.88	0.01	0.76	0.02
	Logistic Regression	0.91	0.01	0.91	0.01	0.83	0.02
	Support Vector Machine	0.91	0.01	0.91	0.01	0.82	0.02
	Neural Network	0.94	0.01	0.94	0.01	0.89	0.02
	Decision Tree	0.95	0.01	0.95	0.01	0.90	0.01

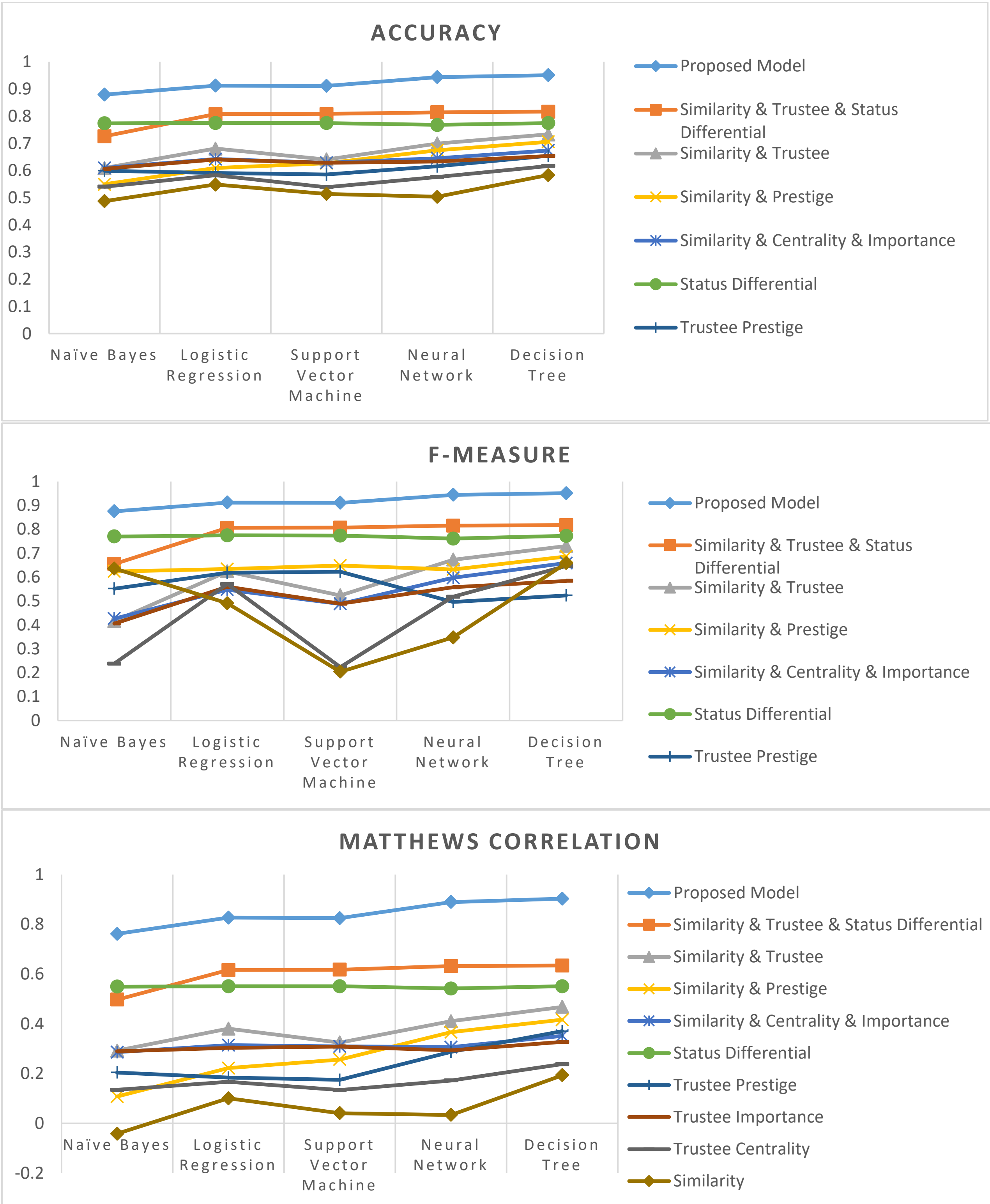

Figure 3.2- Results: Performance of the Proposed Model and Baselines

To test the significance of the effects of feature type and classification method, I ran several ANOVA tests. Particularly, I ran ten one-way ANOVAs to examine the effect of the feature type for each classification method and six one-way ANOVAs to examine the effect of the classification method for each feature type, and for all three performance measures. I also ran related pairwise post-hoc tests (Tukey, Scheffe, Bonferroni).

The results of the ANOVA test show that the effect of the feature type was statistically significant ($p<.001$) for all three performance measures and for all classification methods. All the pairwise post-hoc tests were significant ($p< .001$), implying that the predictive performance of feature types was significantly different from one another for all the classification methods. Furthermore, the main effect of the classification method was statistically significant ($p<.001$) for all three performance measures and for all feature types. Overall, the evaluation results provide evidence of superior predictive performance of the proposed model, compared to previously used models.

3.7. Discussion

Previous methods for trust prediction in signed social networks have largely overlooked trustor/trustee behavioral characteristics. Previous trust prediction models are computational models that consider trust prediction as a special case of sign prediction. They used existing users' data, such as trust voting history (e.g., "web of trust" or "block list"), to predict future trust/distrust relationships. These types of models lack the underlying theoretical foundation of user trusting behaviors and merely focus on computational performance. In addition, most previous models just focused on very limited dimensions of trust/distrust behaviors, such as similarity between users.

Based on socio-psychological trust theories, there are social and behavioral related predictors in trust/distrust relationships, which have been ignored in previous studies. The trustor behavioral and social characteristics have important roles in trusting behaviors regardless of the trustee characteristics. I particularly defined two predictors based on the trustor characteristics: trustor status and trustor trust propensity. Moreover, previous research used very limited network structural characteristics of the trustee, such as trustee centrality or PageRank importance as the trustee popularity or prestige interchangeably. They completely ignored the important characteristic of signed social networks. In signed social networks, popularity (importance) and prestige are two distinct concepts and need to be defined and measured in different ways.

Popularity is the importance or centrality of a node and can be used as structural assurance for trusting behavior, while prestige counts for the net value position (net trust value) of a node. Thus, high popularity means that the trustee is a well-connected and well-known user, but does not necessarily mean that the trustee is a trustworthy user (the number of received trust links is greater than the number of distrust links). I distinguished between trustee structural importance and trustee prestige and defined a set of related features for each of the predictors. Furthermore, I redefined all the behavioral and situational dimensions of trust and specified them for the context of online social networks.

Last but not the least, since distrust information usually is unavailable, previous trust prediction models predict only trust links (and not distrust links). Distrust information is usually more valuable than trust, since users are more confident when expressing their distrust. I used both trust and distrust information to predict trust/distrust links.

I believe that focusing only on limited dimensions of trusting behavior, using only limited structural information, and ignoring the socio-psychological theories of trust result in an incomplete understanding of trust formation and inaccurate trust prediction in online social networks. I addressed the gaps in the existing literature and proposed a comprehensive set of structural and behavioral predictors based on the relevant socio-psychology and computational network theories. Based on that, I proposed a novel theory-based supervised predictive model to predict trust/distrust links in social networks. Empirical evaluation results show the superior fit and predictive performance of the proposed predictive model over previous models that used only limited structural dimensions of online social networks.

There are certainly limitations that need to be addressed in future research. I examined the trust prediction problem as a binary classification problem, classifying a potential link as either a trust or distrust link. Further research is needed to synthesize link prediction (i.e., predicting the likelihood of a potential link) and trust prediction. The evaluation was based on a single dataset and used five standard classification methods. A more comprehensive evaluation may use more datasets from different sources and more classification methods to test the generalizability of my findings. Furthermore, I used a single snapshot of data and applied cross-validation to estimate prediction performance. Future research may acquire longitudinal data and run more realistic evaluation, i.e., training on older data and testing on newer data.

CHAPTER 4

Essay 3: Lexicon-Based Trust Mining in Online Social Networks – A Feature-Based Text Mining Model to Mine Trust Reviews in Online Social Networks

4.1. Introduction

Social media has become an important part of people's personal and professional life (or today's society). Social media is a group of Internet-based applications that build on the foundations of Web 2.0, which enables the creation and exchanges of user-generated content (Kaplan and Haenlein 2010). Social media has had significant influence in the way people connect, communicate, and interact, resulting in an effective form of virtual collaboration, content production, and knowledge sharing. Trust, as a fundamental element of any human social relationships, is even more critical, to have such effective online interactions. In the virtual world, computer-mediated communications replace face-to-face contact. Technology-mediated communications create challenges for effective social interactions. There are social cue deficiencies in computer-mediated interactions since body language and physical surroundings cannot be easily realized through computer channels (Ma and Agarwal 2007). Thus, developing new relationships, sharing personal information, or trusting published information or reviews, can be a more difficult decision compared to face-to-face communications.

Effective governance mechanisms are needed to maintain effective online interactions and sustainable functionality of online communities. Many communities have developed a self-governing mechanism in which the community selects a small group of trusted members to help the community work effectively. This requires the community to decide on the trustworthiness of the candidates, which is usually done by a peer evaluation process. For example, Wikipedia, one of the most successful and effective online communities, has a group of trusted members (Admins) who help to maintain the community functionality.

Wikipedia is the fifth most-visited website, which produces high-quality peer-produced content using the open source online collaboration model. Previous research has studied the Wikipedia online collaboration model and its governance mechanisms (Faraj et al. 2011; Kittur et al. 2007; Ransbotham and Kane; 2011). Some of the previous studies focused on Wikipedia from the perspective of an online social network (Kane 2009; Kane et al. 2014; Ransbotham et al. 2012). The role of Wikipedia trusted Admins is critical to its success. Hence, it is very important to understand the characteristics of such trusted members and the important factors that influence trusting decisions in the context of an online social network.

Previous research in face-to-face communications has shown that perceived trustworthiness of a party is a key predictor for trusting/distrusting that party (Gefen et al. 2003; Mayer and Davis 1999; Mayer et al. 1995; McKnight et al. 1998). In line with previous findings, I also believe that perceived trustworthiness of a trustee candidate is one of the most important determinants of trust/distrust decisions in online communities. Perceived trustworthiness is affected by the qualities that help to mitigate the concerns in trust relationships. Previous studies have identified characteristics, such as ability, reliability, honesty, and altruism as trustworthiness qualities, that help to create a positive sentiment or expectation about the person (Mcknight et al. 1998). The most agreed upon characteristics of trustworthiness are benevolence, competence, integrity

(Butler 1991; Gabarro 1978; Mayer et al. 1995; Mcknight et al. 1998), and predictability (McKnight et al. 2002; Mcknight et al. 1998).

Despite recent studies in online social networks, we still know little about the characteristics of trusted members in online communities. That is what motivated me to study the trustworthiness characteristics that are important in trusting members of online communities. User-generated content (UGC) is the main building block of online social interactions. Valuable insights can be gained from user-generated content by using text mining techniques. To gain insights into what the actual important trustworthiness characteristics are for members of online social networks in their trusting decisions, I mine the trust review texts from a real-world dataset.

I believe that trust review texts reflect the actual viewpoints of online community members. Furthermore, findings from the related behavioral research on trustworthiness in face-to-face communications should be considered and adapted to the online context. This study makes two main contributions. First, I build domain-specific trustworthiness lexicons for online social networks based on related behavioral foundations and text mining techniques. To the best of my knowledge, this is the first trust lexicon developed in the context of online social interactions. Second, I propose a lexicon-based text mining system that automatically extracts important trustworthiness characteristics from user-generated content (trust reviews). The empirical evaluations show the superior performance of the proposed text mining system over the baselines.

4.2. Related Work

Trust has been well studied in face-to-face communications. Trustworthiness beliefs have been found to be a key factor in trusting others (Gefen et al. 2003; Mayer and Davis 1999; Mayer et al. 1995; McKnight et al. 1998).

Early trust theories assumed that trust grows gradually by individuals' interactions over time (Blau 1964; Rempel et al. 1985). Accordingly, when individuals first meet, there is a small initial trust between them. However, several empirical studies showed that there might be a high level of initial trust at the first time two individuals meet. McKnight et al. (1998) proposed a model of initial trust formation addressing this paradox. Initial trust is influenced by trusting beliefs (trustworthiness beliefs), trust propensity, and situational factors. Trust propensity is an inherent characteristic of a trustor's general willingness to trust others. Situational factors are the favorable and protective structures that facilitate the success. Trustworthiness beliefs are the only factor that is directly related to a trustee's qualities. Trustworthiness qualities have been conceptualized as a set of characteristics including benevolence, competence, integrity, and predictability (Mayer and Davis 1999; Mayer et al. 1995; McKnight et al. 1998).

Benevolence

Benevolence is a trustor's belief about a trustee's goodwill toward the trustor, that the trustee cares about the trustor and act in the trustor's interests (Mayer et al. 1995; McKnight et al. 1998). A trustee is benevolent to the degree of her/his openness, loyalty, concern, support and help, acting in the other party's best interest, and caring about others well-being (Colquitt et al. 2007; Mayer and Davis 1999; Mayer et al. 1995; McKnight et al. 2002). Benevolence basically relies more on personal experience or interactions between the trustor and trustee. A trustee's benevolence has been measured by: being concerned about the trustor's welfare, caring about the trustor's needs and desires, not knowingly doing anything to hurt the trustor, looking out for what is important to the trustor, and helping the trustor (Colquitt et al. 2007; Mayer and Davis 1999).

Competence

Competence is a trustor's beliefs about a trustee's ability, skills, and power to do what is needed by the trustor in a specific domain (Mayer et al. 1995; McKnight et al. 1998). Competence is not the trustee's general capabilities or knowledge, but rather, it is domain specific. A trustee is competent as she/he has enough capability, proficiency, ability, expertise, knowledge, and talent in a specific context and performs her/his role very well and in an effective way (Colquitt et al. 2007; Mayer and Davis 1999; McKnight et al. 2002). A trustee's competence has been measured by the capability of performing the job, being successful at doing it, having much knowledge about it, having specialized capabilities that can increase performance, and being well qualified (Colquitt et al. 2007; Mayer and Davis 1999).

Integrity

Integrity is the trustor's belief in a trustee's truthfulness, that the trustee adheres to sound moral and ethical principles, makes good faith agreements, tells the truth, and fulfills promises (Mayer et al. 1995; McKnight et al. 1998). A trustee's integrity is the extent of her/his truthfulness, honesty, fairness, promise keeping, commitment, sincerity and authenticity, bias suppression, ethicality of decision making, and credibility (Colquitt et al. 2007; Mayer and Davis 1999;

Mayer et al. 1995; McKnight et al. 2002). Unlike benevolence, integrity is based more on the trustee's characteristics, than on the interactions or relationships between the trustee and a trustor (McKnight and Chervany 2001). A trustee's integrity has been measured by having a strong sense of justice, sticking to her/his word, being fair in dealings with others, having consistent actions and behaviors, and behaving based on sound principles and values (Colquitt et al. 2007; Mayer and Davis 1999).

Predictability

Predictability is the trustor's belief that a trustee's actions (good or bad) are sufficiently consistent so that the trustor can predict the trustee's future actions in a given situation (McKnight et al. 1998; McKnight and Chervany 2001). Predictability could mean that the trustee always acts in the same way, such as consistently meeting the trustor's preferences. A trustee's predictability is the extent of her/his consistency of actions (McKnight et al. 1998).

4.2.1. Trustworthiness in Online Communities

In addition to technological requirements, effective governance mechanisms are required to have sustainable functionality of online communities. Community leaders, as one of the effective governance mechanisms, play a critical role in the sustainability of online communities' function (Preece 2000). Community leaders have an important role in generating community participation, building relationships, and, particularly, promoting collaboration and trust among community members (Koh et al. 2007; Preece 2000). Many communities have developed a self-governing mechanism in which the community selects a group of trusted leaders. In such online communities, leaders govern the community, set policies and procedures, and help other members to work effectively. This is an effective strategy since it helps to reach consensus decisions more easily (Kittur and Kraut 2010). For example, Wikipedia, as a collaborative content creation project, enables multiple users to constantly improve this free encyclopedia. Previous research has found that the quality and accuracy of Wikipedia articles are comparable to those of established sources such as Britannica Encyclopedia (Giles 2005). One of the key reasons for such high quality articles is the peer-review rating system that Wikipedia uses to rate its content. The peer-review system is based on the collaboration of Wikipedia editors to work with one another in modifying the contributions made by others (Kittur et al. 2007).

Wikipedia editors can get promoted to administrators (admins), who help in community maintenance. Administrators (admins) are trusted users who have access to additional technical features, and their actions can impact the entire community. Granting administrator status is considered to require a high level of trust from the community. Hence, selecting the trusted members is done by a voting process in which Wikipedia community votes in favor of or against the promotion of the admin candidates. Many voters decide to trust or distrust a candidate based on a set of qualities, called trustworthiness characteristics. Voters judge the candidates' trustworthiness based on the candidates' past work, behavior, and also any previous interactions with them. Similar to product review communities, voters can also read other members' reviews about the candidates for more information. There must be certain qualities that members of such online collaborative communities seek in candidates to trust them as admins.

Researchers have studied the social aspects and network characteristics of Wikipedia with respect to its content quality (Kane 2009; Kane et al. 2014; Ransbotham et al. 2012). They viewed Wikipedia as a content–contributor network and studied it using a two-mode social network analysis. Kane (2009) found that an article position in the content-contributor network (degree centrality and eigenvector centrality) is related to the article quality. Ransbotham et al. (2012) found that locally central content (greater intensity of work by contributors) and globally central (shorter paths to the other collaborative content in the network) content generate higher viewership.

As discussed above, Wikipedia admins play a critical role in community functionality and quality of its peer-produced content. However, the RfA process (admin selection process) **__** how members trust others and what qualities they seek in candidates to trust them as network admins **__** has not been studied in detail yet. This study follows the network view of Wikipedia and analyzes the trustworthiness qualities that are important for members to trust others as network admins.

4.2.2. Feature-Based Opinion Mining

Opinion is a key predictor of human decisions and behaviors (Liu 2012). Public opinion is important for both individuals and businesses in decision making. Business examples can be predicting the success of a new product, popular model of a product, causes of not selling well, and future niche markets. After Web 2.0, online public opinion has become more critical for business success. With the popularity of user-generated content (UGC), opinion mining has attracted enormous attention in recent years. UGC refers to any form of content created and shared by online users, such as blogs, wikis, tweets, posts, reviews, comments, podcasts, images, audios, videos, and podcasts.

A vast amount of user-generated content is generated every second by millions of social media users, resulting in so-called "Big Data". Most of the user-generated content is unstructured textual contents, creating a huge rich repository of analytic information of human behavior and interactions. However, the noisy and dynamic nature of unstructured textual data makes it a big challenge to extract useful information and valuable insights. Recent advances in text mining and access to the UGC – a huge and rich repository of documented opinions – make opinion mining a very attractive topic. Many computational models and machine learning systems have been built to collect and effectively analyze the UGC to understand and extract patterns, associations, and insightful knowledge.

Opinion mining, as one of the most useful applications of text mining, attempts to identify and extract human opinion and/or polarity (sentiment) from a text written in natural language. It uses multiple techniques, such as natural language processing, computational linguistics, and text analytics. Liu (2012) defined opinion mining (also called sentiment analysis) as the task of analyzing people's opinions, attitudes, sentiments, evaluations, and emotions toward an item – such as a product, service, organization, individual, or topic – and its attributes. Many automatic systems have been developed to collect, examine, and extract public opinions from customers' reviews about products and services in online retailing websites (such as OPINE by Popescu and Etzioni (2007) and Red Opal by Scaffidi et al. (2007)). An ideal opinion mining tool is supposed to process a set of textual data for a given item (product), extracting a list of item (product) features and aggregated opinions on each of them (Dave et al. 2003).

Feature (aspect/attribute) based opinion summarization is one of the important forms of opinion mining (Dave et al. 2003; Liu 2012). The key task is to find features (attributes) of the item under review and to present an aggregation of opinions about each of those features. For example, the following review expresses two positive opinions about two features (screen and battery life) and one negative opinion about one feature (camera) of a cellphone: “the screen is very bright and high-contrast, battery life is long, however, the camera resolution is very poor”.

There are two main tasks in feature-based opinion mining: feature extraction and semantic orientation (sentiment) classification. Feature (attribute) extraction is a specific form of the general information extraction problem (Liu 2012), which finds the topic or attribute that is evaluated in the text. Semantic orientation (sentiment) classification identifies the attitude or sentiment toward each of the extracted features.

Feature-based opinion mining has been extensively studied in previous research (Ding et al. 2008; Ghani et al. 2006; Hu and Liu 2004; Jin and Ho 2009; Kovelamudi et al. 2011; Popescu and Etzioni 2007; Scaffidi et al. 2007; Zhang et al. 2010). The two main approaches used in previous studies are the supervised learning approach and the unsupervised (lexicon-based) approach. Furthermore, topic modeling has been found to be unacceptable in feature-based opinion mining. I discuss each approach in the following section.

Topic Modeling

Topic modeling is one of the special applications of text mining to identify latent semantic structures (topics) in documents. The idea is that particular words about a topic appear together. Thus, by finding semantic correlations among words, hidden topics can be extracted. Topic models are usually unsupervised probabilistic models that generate multiple sets of similar word clusters that can be named to multiple topics. One of the most effective methods in topic modeling is Latent Dirichlet Allocation (LDA) by Blei et al. (2003), which is a three-level hierarchical Bayesian model.

Topics extracted from topic modeling might be used as the opinion features in feature-based opinion mining. However, it has been shown that global topic models, such as LDA (Blei et al. 2003), are not suitable for identifying opinion features (Titov and McDonald 2008). Liu (2012) believe that the reason is behind the basics of LDA algorithm. LDA identifies topics and word probability distribution in each topic, based on the topic distribution differences and word co-occurrences among documents. However, opinion documents, such as reviews, are similar because they are about a particular type of product or service and talk about the same product features. Hence, they are all about a common topic, meaning that no topic distribution differences exist. This makes topic models ineffective in feature-based opinion mining.

Supervised Learning Approach

Supervised learning methods seek and learn patterns in data and require a manually annotated training dataset to train the predictive model (classifier). Hu and Liu (2006) used language patterns and proposed a supervised sequential pattern mining method to identify opinion features from Pros and Cons reviews. The method basically seeks class sequential rules that have product features as class items. Jin and Ho (2009) proposed a supervised model, using lexicalized Hidden Markov Models (HMMs), that learns patterns from training data to extract product features. The model integrates linguistic features, such as part-of-speech and lexical patterns, into HMMs. Kovelamudi et al. (2011) proposed a supervised domain-independent model to extract product attributes from user reviews. Ghani et al. (2006) used supervised and semi-supervised learning techniques to the product descriptions on retailer websites for product attributes extraction.

There are some problems with supervised learning methods that make them less attractive for feature-based opinion mining (Liu 2012). First, supervised learning methods require a large manually annotated training dataset to achieve acceptable performance. Creating a large annotated training dataset is expensive and inefficient, especially in social media text mining. Second, supervised learning methods are mostly suitable for document-level, rather than sentence-level opinion mining, as documents are long and contain more features for classification. That makes them ineffective for social media text mining, because user-generated content, such as tweets and comments, tends to be very short. Third, the performance of a supervised learning method can vary a lot in different domains. To address that, domain adaptation of classifiers has been proposed but its effectiveness is still under debate.

Unsupervised (Lexicon-Based) Approach

Unsupervised techniques are more domain-independent and much cheaper to develop and modify (Carenini et al. 2005). Hence, it has attracted more attention in feature-based opinion mining research. Most unsupervised models in feature extractions and sentiment prediction are lexicon-based methods (Ding et al. 2008; Hu and Liu 2004; Popescu and Etzioni 2007; Scaffidi et al. 2007; Zhang et al. 2010). Lexicon-based methods have shown to perform quite well in many different domains (Liu 2012). For the feature extraction task, they usually use a frequency-based method to extract the most frequent terms as possible candidates for features. For the sentiment prediction task, they usually use a sentiment lexicon (a list of opinion words) to determine the semantic orientation toward each extracted feature.

The work by Hu and Liu (2004) is among the first attempts to study the problem of generating feature-based summaries of customer reviews. They proposed quite an effective lexicon-based method using opinion-bearing words and association mining to extract product features. The proposed model seeks frequent item sets, a group of words that occur together, by association mining. The idea is that frequent words or phrases that are talked about by many customers are likely to be product features. Semantic orientations of features are predicted using an opinion lexicon obtained from WordNet (Miller 1995). WordNet synonyms and antonyms were used to produce the opinion lexicon, based on the fact that terms share the same/opposite orientation as their synonyms/antonyms.

Following Hu and Liu (2004), several other lexicon frequency-based methods have been proposed for feature-based opinion mining. Popescu and Etzioni (2007) proposed OPINE, a review-mining system, which improved the precision of the feature extraction task over that of

Hu and Liu (2004). OPINE identifies product features by computing Pointwise Mutual Information (PMI) between the candidate phrase and a group of discriminator phrases. For example, discriminators for “scanner” are: “of scanner”, “scanner has”, “scanner comes with”, etc. A limitation of their proposed method is that the product class (e.g., camera) must be known. Furthermore, they estimated PMI from web search engine hit counts (Turney 2002), which is a time-consuming method.

Ding et al. (2008) further improved the work of Hu and Liu (2004) and proposed a lexicon-based model, called Opinion Observer, for opinion feature extraction. They defined a set of linguistic rules to address context-dependent words, special opinion words, conflicting opinion words, and phrases in a sentence. In another related study by Scaffidi et al. (2007), an automatic system called Red Opal was proposed, which enables users to find a product based on the desired product feature. The system examines customer reviews, identifies product features, and scores each product on each feature. Users can then rapidly locate the product with the specific feature they are looking for. The main idea behind feature extraction is that some words occur far more frequently in the review text than in a random section of English text of equal length. Thus, the frequency of a candidate term in the review is compared to its frequency in a general text to determine the true features. Zhang et al. (2010) proposed a double propagation method to extract and rank features based on feature frequency and feature relevance. The proposed method uses an initial seed opinion lexicon and the web page ranking algorithm HITS. The basic idea in double propagation is that features (nouns/noun phrases) are usually associated with opinion words (adjectives). Thus, one can be recognized if the other is known. Based on that, the propagation or bootstrapping process continues until no more opinion words or features can be found.

The studies discussed above show that the unsupervised (lexicon-based) approach is the most suitable for feature-based opinion mining of user-generated content. The key element of lexicon-based methods is the opinion lexicon. In the following section, I discuss methods for lexicon development.

Lexicon Development

An opinion lexicon is a set of words or phrases that are used to express a positive or negative sentiment. There are three main approaches that have been proposed to build a lexicon (Liu 2012): manual approach, dictionary-based approach, and corpus-based approach. The manual approach is inefficient and costly, and it is rarely used. The dictionary-based approach is easier than the corpus-based approach. The reason is that various dictionaries are available that could be used to generate the initial seed words of the lexicon. In the dictionary-based approach, a small set of seed words are manually created, then the lexicon is expanded by searching for their synonyms and antonyms from a dictionary such as WordNet (Miller 1995). The searching and adding step is an iterative process that will continue until no new words are found. At the end, a manual inspection is usually done to clean up the lexicon. The dictionary-based approach was first used by Hu and Liu (2004) in feature-based opinion mining and later in multiple related studies (Blair-Goldensohn et al. 2008; Esuli and Sebastiani 2005; Kim and Hovy 2004; Peng and Park 2011).

Although the dictionary-based approach is very easy and fast for creating a large lexicon, it produces many irrelevant words. That is the reason why a manual inspection is required to clean up the automatically generated lexicon. This makes it indeed a costly and time-consuming approach. But the main problem is that the final lexicon cannot be used to predict the semantic orientation of context or domain-related words. The dictionary-based method seeks synonyms and antonyms from a general and domain-independent dictionary to create the final lexicon. The truth is that in many applications, words have context (domain) dependent meanings that need to be considered in opinion mining.

To address the mentioned problems, the corpus-based approach has been proposed. The main idea is to expand a general lexicon (e.g., one generated from a dictionary) with new words from a domain corpus. One early approach was to use a set of seed words along with a set of linguistic rules to find additional words from the corpus (Hatzivassiloglou and McKeown 1997). For example, if an unknown word in the corpus is connected with a known word from seeds with "AND", the new word from the corpus would have the same semantic orientation and can be added to the lexicon. However, it has been shown that it is possible that a word has multiple semantic orientations in the same domain (Ding et al. 2008). For example, the word "long" in the following two sentences has two opposite sentiments: a positive sentiment in "the battery life is very long" and a negative sentiment in "it takes a long time to focus". Ding et al. (2008) proposed multiple linguistics rules, such as conjunctions rules, to consider both the feature and the opinion word, rather than only the opinion word.

Recent studies considered using both domain information from the corpus and a dictionary-based general lexicon. Peng and Park (2011) proposed a sentiment dictionary generation method, Constrained Symmetric Nonnegative Matrix Factorization (CSNMF), using both WordNet and a large social media corpus. They showed that the generated sentiment dictionary by combining dictionary and corpus outperform a method using only one of them.

4.3. Proposed Model

In this study, I propose a lexicon-based text mining system to identify, extract, and classify trustworthiness characteristics from user-generated content (trust reviews). The proposed system has two main parts: trustworthiness lexicon development and trustworthiness characteristic classification.

I adopt the lexicon-based approach because of its relative advantages compared to supervised learning and topic modeling. As discussed earlier, there are multiple problems with supervised learning methods that make them ineffective for feature-based opinion mining. Creating a large manually annotated training dataset and the need for large documents with enough textual features are among the mentioned problems. The final goal of the proposed text mining system is to extract and classify trustworthiness characteristics from online trust reviews. Most of the reviews are short and every single mentioned characteristic has to be extracted at the sentence level. Thus, supervised learning methods will not serve this purpose.

Topic modeling is also not a suitable method for feature-based opinion mining, as discussed earlier. Previous studies showed that global topic models, such as LDA, do not work for identifying opinion features (Titov and McDonald 2008). The reason is the basic assumption behind the topic models: the differences in topic distribution. However, opinion documents, such as reviews, are similar because they are about a particular type of product or service. Hence, they are all about the same topic, meaning that no topic distribution differences exist. Thus, it is almost impossible to extract different features (topics) from opinion documents using topic models.

In this study, I empirically show the ineffectiveness of topic modeling in feature-based opinion mining. I applied LDA on a real dataset from Wikipedia RfA (admin election) reviews. If LDA works well, it should identify the different topics (or trustworthiness character tics) that the voters have discussed in their reviews. But the topics extracted from LDA results in my empirical study were not meaningful. As discussed earlier, LDA cannot extract meaningful features (trustworthiness characteristics), since all the reviews talk about the same topic (trusting the candidates).

4.3.1. Trustworthiness Lexicon Development

Lexicon-based text mining systems require lexicons. There are not any previously created lexicons in the trust domain. Therefore, the first step in building the proposed lexicon-based trustworthiness mining system is to develop trustworthiness lexicons. I generate initial seeds for the trustworthiness lexicons based on previous behavioral trust studies in the context of face-to-face communications. Next, I expand and adapt the lexicons to the domain of online social communities. I use both dictionary-based and corpus-based methods to develop trustworthiness lexicons.

The proposed method for trustworthiness lexicon development is summarized as the following steps (see figure 4.1):

1- Build initial trustworthiness lexicons based on the behavioral literature in trust
2- Expand seed lexicons using a dictionary-based method
3- Adapt the dictionary-expanded lexicons to the domain of online social communities using a domain corpus:
    a. Perform preprocessing
    b. Identify candidate features
    c. Filter and prune candidates

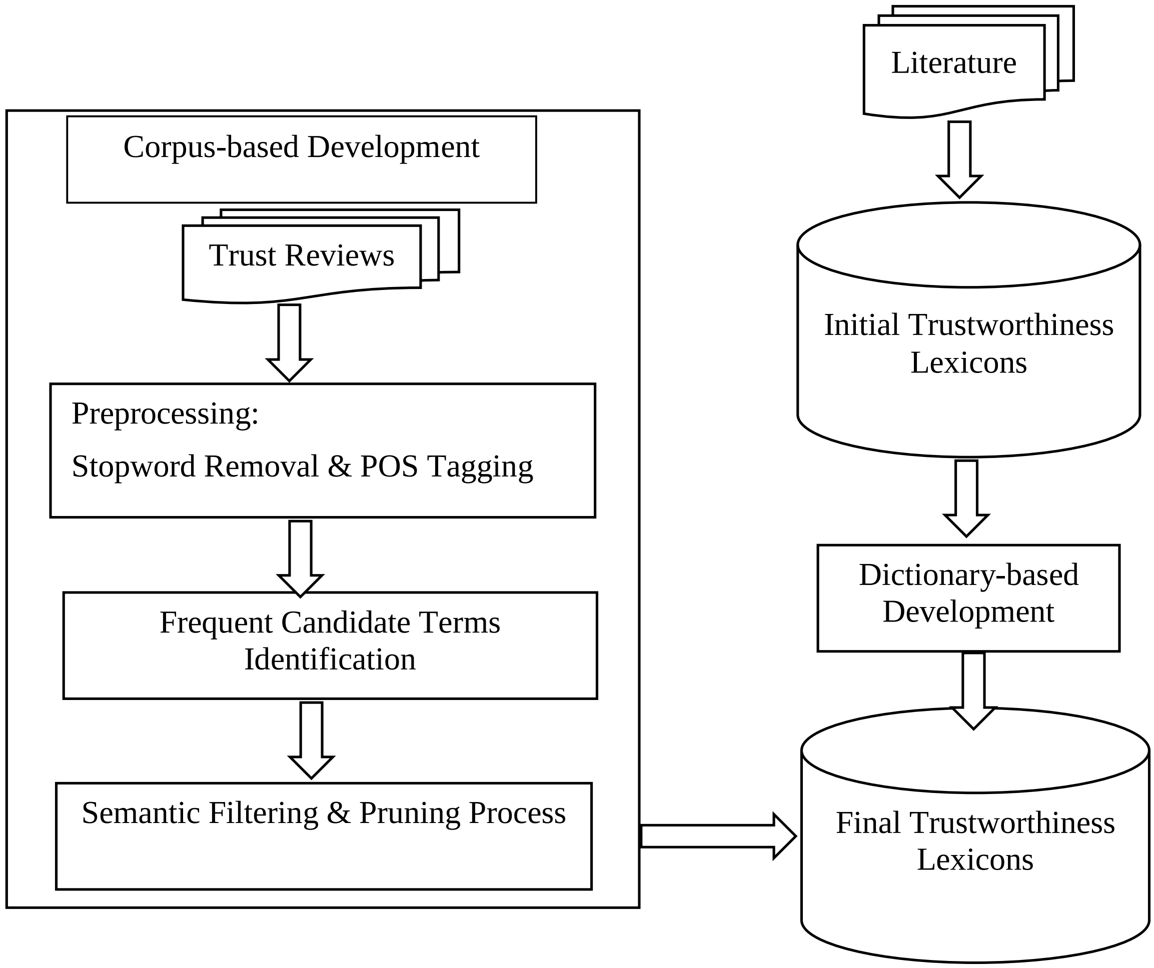

Figure 4.1- The Proposed Method for Trustworthiness Lexicon Development

For the purpose of generating initial seeds, I first identify trustworthiness seed words rooted in earlier behavioral theories of trust. Previous research has identified four main dimensions for trustworthiness characteristics: benevolence, competence, integrity, and predictability (Mayer et al. 1995; Mayer and Davis 1999; McKnight et al. 1998; McKnight et al. 2002). Accordingly, I create four initial lexicons.

Initial seed words were selected using measure items from previous research (McKnight et al. 1998; McKnight et al. 2002). The four trustworthiness lexicons are expanded using both a dictionary-based method and domain information. WordNet has been used in many previous dictionary-based lexicon development methods with valid results (Hu and Liu 2004; Popescu and Etzioni 2007; Ding et al. 2008). Hence, I use WordNet to expand the initial seed terms. In contrast to sentiment lexicons, trustworthiness lexicons should include both positive and negative terms related to trustworthiness qualities. For example, for “competence” characteristic, we are interested in both “experienced” and “inexperienced” words. Therefore, both synonyms and antonyms from WordNet are added to the initial seed words.

The dictionary-based trustworthiness lexicons are further expanded by incorporating domain-dependent opinion words. I use a domain corpus to find the opinion words and phrases that are specific to the domain of trust in online social networks. The first step in corpus-based methods is pre-processing the corpus, which consists of stop word removal and POS tagging. Stop words are the most common words, usually articles, prepositions, and pronouns (e.g., I, she, this, and those) that do not give any specific meaning to the corpus. Therefore, all the English stop words, numbers, punctuations, symbols, in addition to any meaningless words and phrases, such as web address, links, and http tags, or non-English words, are removed from the corpus. Then, all the whitespaces created in the stop word removal step are stripped. All the extra spaces are shortened to one space.

The next step is part of speech (POS) tagging, which enables syntactical analysis of the corpus. POS tagging annotates each word (token) into the relevant part of speech, such as adjectives, adverbs, nouns, and verbs. This step could be very important for this study. For example, in sentiment analysis, most of the sentiment words are adjectives, while in product feature extraction, most of the product features are nouns. Trustworthiness characteristics can be described in reviews as an adjective (competent), an adverb (consistently), a noun (knowledge), or a verb (helps). A sentence example can be: “he is very responsible”, “he acts maturely”, “he has enough expertise”, or “he cares about others”. Thus, I keep all the adjectives, adverbs, nouns, and verbs, and discard all the rest, such as coordinating conjunctions, determiners, modals, pronouns, and Wh determiners.

After the preprocessing step, all candidate terms that have the potential to be added to trustworthiness lexicons must be identified from the corpus. Frequent words and phrases are used to find potential features (Hu and Liu 2004; Popescu and Etzioni 2007; Zhang et al. 2010). Hence, the term frequency matrix is built to find the frequencies of all terms in the corpus. I use the document-term matrix (dtm), where cell(i,j) represents the frequency of term j in document i. Candidate frequent features are extracted with frequency greater than an experimentally set threshold (Popescu and Etzioni 2007).

After selecting the frequent terms, a filtering process is needed to select the final terms to be added to the lexicons. Each identified frequent term needs to be evaluated for addition to the relevant trustworthiness lexicons or to be discard. First, all the synonyms of a frequent term are found from WordNet. If any of the synonyms matches any of the lexicon terms, the frequent term will be added to that lexicon. The algorithm also adds only unique terms to the lexicons. If the frequent term is already among the lexicon terms, it will not be added to the lexicon. Thus, the lexicons do not contain any duplicates. Consequently, the heavy load of manual inspections of generated lexicons is reduced.

4.3.2. Trustworthiness Extraction and Classification

The proposed lexicons can be used for mining online trust reviews. Using a lexicon-based text mining system, trustworthiness characteristics mentioned in trust reviews are extracted and classified to the related trustworthiness classes. The proposed method to extract and classify the trustworthiness characteristics from trust reviews is summarized as the following steps (see figure 4.2):

1- Preprocess the review
2- Identify the candidate terms

3- Determine the average information association (PMI & Co-Occurrence) of each candidate term with the trustworthiness lexicons

4- Assign the candidate term to the relevant trustworthiness lexicon

5- Evaluate

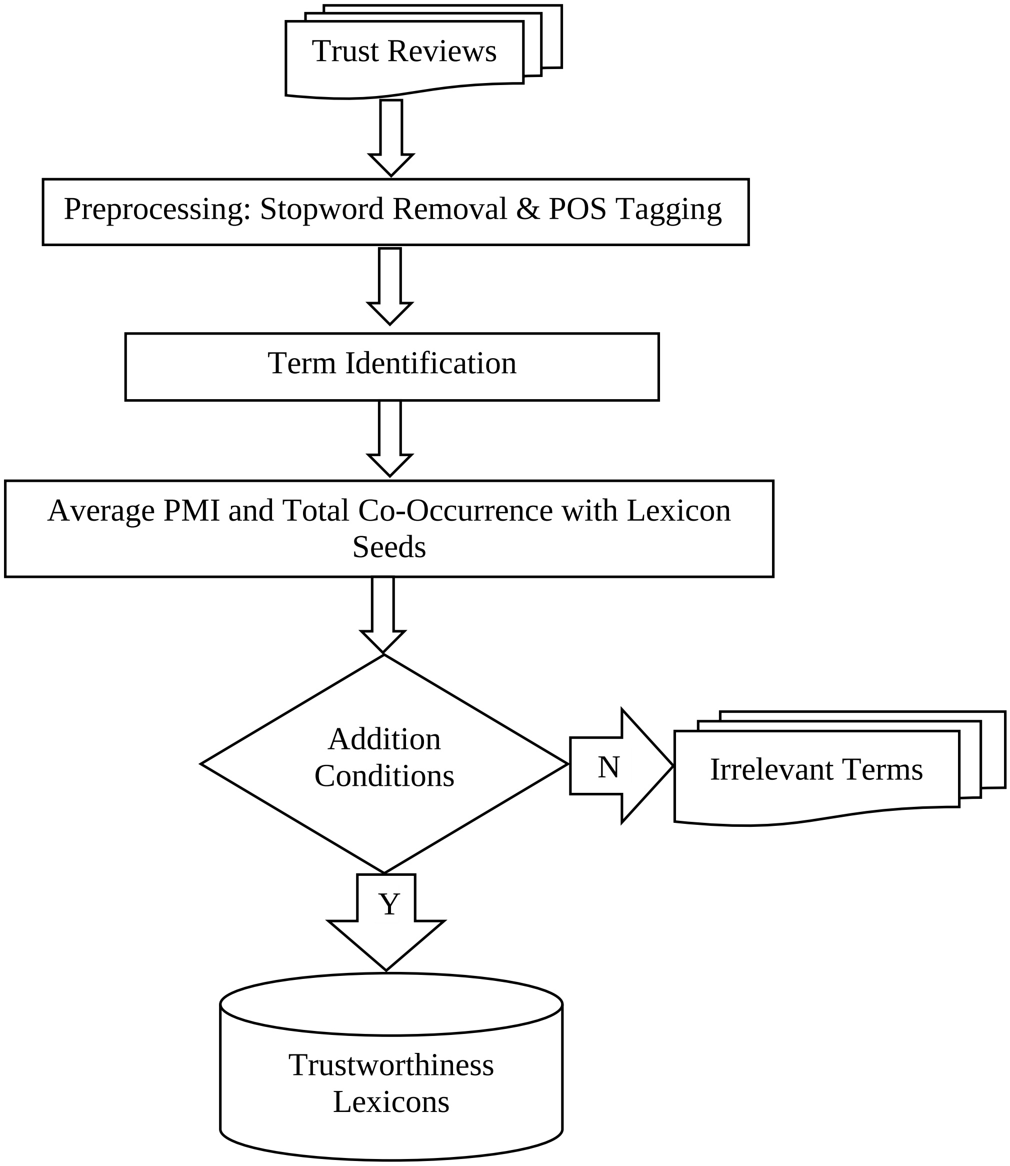

Figure 4.2- The Proposed Method for Extracting and Classifying Trustworthiness Characteristics from Trust Reviews

First, each review is preprocessed and the relevant document-term matrix (dtm) is built. I use a Pointwise Mutual Information (PMI) based method for term classification. PMI has been used in previous research on sentiment analysis and feature-based opinion mining (Turney 2002;

Popescu and Etzioni 2007). PMI, by Church et al. (1991), is a statistical measure of information association, which was primarily applied to collocation analysis. The main idea of the method is that co-occurrence of words is a measure that they convey mutual information. The formula is defined as the following: PMI $(\mathrm{term1}\,,\mathrm{term2}) = \log 2(\frac{\Pr(\mathrm{term1} \wedge \mathrm{term2})}{\Pr(term1) * \Pr(\mathrm{term2})}$, where $\Pr(\mathrm{term1} \wedge \mathrm{term2})$ is the joint probability of term 1 and term 2 occurring together in the document, and is calculated by dividing the number of their co-occurrences by the total number of words in the document.

For each lexicon, the co-occurrence and PMI scores of the review terms with all the seeds in the lexicons are calculated. The total co-occurrence and the average PMI with all seeds are calculated.

To classify a term into the related lexicon, three criteria need to be met. First, to be considered as a candidate for classification, the term's total co-occurrence with the lexicon seeds must be greater than a defined threshold. This condition is set for the cases that co-occurrence with one lexicon is very small and for all other lexicons is zero. For example, a term's total co-occurrence with all terms in lexicon A is two and the co-occurrence with all other lexicons is zero. In this case, if we only consider maximum co-occurrence, the term will be classified as belonging to lexicon A, which is misleading. A total co-occurrence of two is very low and does not mean that the term associates with the lexicon meaningfully.

Second, the average PMI between the term and the lexicon seeds is the maximum compared to other lexicons. Third, the total co-occurrence of the term with the lexicon seeds is at least four times greater than that with other lexicons.

4.4. Experiment

I created four initial trustworthiness lexicons for Benevolence, Competence, Integrity, and Predictability, including all the seed words from the measurement items in related literature (Mayer and Davis 1999; Mayer et al. 1995; McKnight et al. 1998 and 2002). Next, all four initial lexicons were expanded by adding synonyms and antonyms using WordNet (Miller 1995). To adapt the dictionary-based lexicons to online social communities, a corpus related to the study domain is needed.

I used Wikipedia RfA (Request for Adminship) election data. Wikipedia, as one of the largest and most popular online encyclopedia, is an online collaborative project. All of its members contribute toward improving this online encyclopedia. Members can connect and communicate with other members, like in any other online social networks. Wikipedia administrators are trusted members who are elected by the community members and are granted additional tools to perform certain actions. Members can request for promotion from editor to admin status. I selected to examine the adminship election data because it meets the requirements for this study. In the RfA process, the community members (as trustors) decide whether to trust or distrust a nominated candidate for adminship (a trustee). In addition to the vote, they also express their opinions about the candidate in the form of a trust review. The data was collected from 2003 (since the adoption of the RfA started) through 2013, containing 11,402 users (voters and votees) and 198,275 votes, in addition to the textual reviews (available at the Stanford Network Analysis Project (SNAP)). All the trust reviews were combined to build the domain corpus.

The corpus was cleaned by removing all the blank reviews, non-English reviews, meaningless reviews, or those that just mentioned only the vote (e.g. support or oppose) and did not give any opinions. The preprocessing was done by removing all the English stop words, numbers, punctuations, and symbols, in addition to web addresses, links, http tags, and usernames. Finally, all the letters were converted to lowercase and all additional spaces were stripped to one space. After the cleaning and preprocessing step, 96,415 trustor opinions, including 1,424,984 words, were obtained.

I used MorphAdorner V2.0 Part of Speech Tagger developed by Northwestern University for part of speech tagging. As explained in the proposed method, I kept all the adjectives, adverbs, nouns, and verbs, and discarded all the rest. The term frequency matrix (dtm) was built to find the term frequencies. Following the method by Popescu and Etzioni (2007), I kept 1% of the most frequent terms, resulting in 9162 unique tokens. I used the obtained corpus terms to adapt the dictionary-based developed lexicons to the domain of online social communities. I also used a set of initial domain extracted seeds to be added to dictionary-based generated lexicons. For all four trustworthiness lexicons, the following algorithm was performed using the R-Studio platform:

```
for i in the lexicon:
    let LSyn be the vector of all synonyms for i found in WordNet
    for j in corpus frequent terms:
        let TSyn be the vector of all synonyms of j from WordNet
        if any match found between LSyn and TSyn AND if j is not already in
        the lexicon, then:
        add j to the lexicon
```

The generated trustworthiness lexicons can be used to extract and classify trustworthiness features of trust reviews. Each review was preprocessed and the document-term matrix (dtm) was created. For all four lexicons, I calculated the average PMI and the total co-occurrence of each review term with the lexicon seeds. To be able to calculate PMI, a reference corpus is needed. The reference corpus must include all the terms that normally appear together in the context. Thus, it must include the domain-specific terms, in addition to the general terms. There was not any previously built reference document for the domain of trust in online social communities. I built the reference corpus using a large general purpose synonym-antonym based dictionary (Ordway 2009), in addition to the domain-specific terms. The domain terms were extracted from the trust reviews of Wikipedia RfA. For all trustworthiness lexicons, the following algorithm was performed using the R-Studio platform:

let *l* be the total number of words in the reference document
let *k* be the total number of words in the lexicon
for *i* in review terms:
  let LCount be the number of occurrences of *i* in the reference corpus
  for *j* in the lexicon:
    let RCount be the number of occurrences of *j* in the reference corpus
    let LRCount(*i,j*) be the number of co-occurrences of *i*&*j* in the reference corpus
    $\mathrm{PMI}(i,j) = \log_2((\mathrm{LRCount} * l)/(\mathrm{LCount} * \mathrm{RCount}))$
    $\mathrm{AVEPMI}(i,j) = \sum_{j=1}^{k} \mathrm{PMI}(i,j)/\ k$
    $\mathrm{SUMLRCount}(i,j) = \sum_{j=1}^{k} \mathrm{LRCount}(i,j)$

For the term(*i*) of a review, the average PMI and total co-occurrence with all four lexicon seeds is calculated. Then, term(*i*) is classified to lexicon(*j*), if SUMLRCount(*i,j*) is greater than a minimum threshold, $\mathrm{AVEPMI}(i,j)$ is the maximum (compared to the other three lexicons), and the $\mathrm{SUMLRCount}(i,j)$ is at least 4 times greater than the other three lexicons.

4.4.1. Evaluation

To evaluate the proposed method, I compared the performance of the proposed trustworthiness lexicons with that of baselines. Since there is not any previous study in trust mining in online social networks, I created multiple lexicons to compare the performance of the proposed method to them. The first set of lexicons was generated using initial seeds based on the measurement items for trustworthiness in the face-to-face communications (McKnight et al. 1998). The second set was generated by expanding the first set (literature-based lexicons) by a dictionary-based method using WordNet. The third set was generated based on both dictionary-expanded literature seeds and a domain corpus. All the relevant terms in the domain corpus were used to adapt the dictionary-expanded lexicons to the context of online communities.

A gold standard model is required to enable evaluating the results from text mining systems. The gold standard model was created by using human annotators. I asked two graduate students to manually annotate a random sample of 1000 trust reviews (10% of the total trust reviews), according to the measurement items. Each annotator identified any trustworthiness characteristics that were mentioned in the review. I asked them to resolve the conflicts between their annotations, if any, to get the final values for all four trustworthiness characteristics.

The Cohen's kappa coefficient (Cohen 1960) was calculated to determine the inter-rater agreements between the two annotators. A kappa coefficient value between 0.81–0.99 is almost a perfect agreement (Landis and Koch 1977). The calculated kappa coefficients for all four trustworthiness characteristics were above 0.81, indicating almost perfect agreement between the two annotators (Table 4.1).

Table 4.1- Cohen's kappa Coefficient for all four Trustworthiness Characteristics

Benevolence		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Agreement	Kappa	.851	.058	8.557	.000
N of Valid Cases		100			

Competence		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Agreement	Kappa	.959	.029	9.595	.000
N of Valid Cases		100			

Integrity		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Agreement	Kappa	.851	.058	8.557	.000
N of Valid Cases		100			

Predictability		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Agreement	Kappa	.814	.062	8.288	.000
N of Valid Cases		100			

4.5. Results

The system automatically extracts trustworthiness characteristics from trust reviews and classifies them into the related trustworthiness classes. Four text mining systems were built using the three sets of baseline lexicons, in addition to the lexicons developed based on the proposed method. The proposed trustworthiness lexicons were developed by both dictionary-based and corpus-based methods using a filtering process on frequent terms of the domain corpus. They were built using dictionary-expanded literature seeds, in addition to the filtered domain corpus seeds. The first set of baseline lexicons (Literature) was developed based on the literature seeds. The second set of baseline lexicons (Literature_Dictionary) used dictionary expanded literature seeds. The third set of baseline lexicons (Literature_Dictionary_FullCorpus) used both dictionary expanded literature seeds, in addition to all the relevant seeds from a domain corpus.

The results were evaluated using multiple performance measures, including accuracy, recall, precision, and F-measure. Accuracy is the proportion of actual cases (true or false) that are correctly predicted, Accuracy= (TP+TN)/N. Precision (confidence or exactness) is the proportion of the predicted true cases that are actually true, Precision=TP/(TP+FP). Recall (sensitivity or completeness) is the proportion of actual true cases that are predicted correctly, Recall=TP/(TP+FN). F-Measure (F1 or F-Score) is the harmonic mean of precision and recall, F=(2×precision×recall)/(precision+recall).

Table 4.2 presents the results of the proposed method compared to baseline lexicons. The results show that by using the proposed lexicons, the best performance was achieved in extracting and classifying the trustworthiness characteristics of trust reviews. In addition, the results confirm the importance of adapting the literature to the context of online communities.

The worst performance was achieved by using the lexicons developed based on the seeds from the traditional face-to-face trust literature. It confirms that using a combination of both the dictionary-based and corpus-based methods leads to better performance than using only the dictionary-based method. However, the corpus should be used cautiously by applying relevant filters.

The lexicons developed by using all the terms from the domain corpus showed poorer performance compared to the proposed lexicons. By not filtering the domain corpus, there might not be any advantage of adding the domain information. Not pruning and filtering the domain corpus can cause noise (by adding irrelevant or no-value terms to the lexicons), which deteriorates the classification performance. For example, in some cases, the performance was the same or even worse than when using the lexicons developed by literature seeds.

Table 4.2- Experiment Results

		Accuracy	Precision	Recall	F-Measure
Benevolence	Literature_Dictionary_FilteredCorpus	0.96	0.91	0.89	0.90
	Literature_Dictionary_FullCorpus	0.89	0.82	0.60	0.70
	Literature_Dictionary	0.89	0.82	0.59	0.69
	Literature	0.87	0.71	0.62	0.66

		Accuracy	Precision	Recall	F-Measure
Competence	Literature_Dictionary_FilteredCorpus	0.81	0.74	0.93	0.83
	Literature_Dictionary_FullCorpus	0.73	0.71	0.74	0.73
	Literature_Dictionary	0.70	0.72	0.61	0.66
	Literature	0.71	0.67	0.80	0.73

		Accuracy	Precision	Recall	F-Measure
Integrity	Literature_Dictionary_FilteredCorpus	0.88	0.89	0.82	0.86
	Literature_Dictionary_FullCorpus	0.83	0.76	0.87	0.81
	Literature_Dictionary	0.68	0.62	0.66	0.64
	Literature	0.84	0.82	0.82	0.82

		Accuracy	Precision	Recall	F-Measure
Predictability	Literature_Dictionary_FilteredCorpus	0.83	0.64	0.81	0.72
	Literature_Dictionary_FullCorpus	0.81	0.61	0.85	0.71
	Literature_Dictionary	0.79	0.59	0.84	0.69
	Literature	0.79	0.58	0.83	0.69

4.6. Discussion

Trustworthiness of a trustee, as a key predictor in trusting decisions, is a critical factor for the trustee to be trusted by others. With recent growth of social media, it is important to gain insights into trust formation and trusting behaviors in the context of online social interactions. User-generated content is a unique valuable source of information about online users' opinions. This is the first study in the field of online trust that seeks to understand user trusting behaviors based on user-generated content. In this study, I proposed a lexicon-based text mining method to identify, extract, and classify trustworthiness characteristics from user-generated content (trust reviews).

There are not any previous trust lexicons to be used in studying trust in online social communities. In this study, trustworthiness lexicons for trust in online communities were created based on the related theoretical background and a domain corpus. Domain-specific characteristics were used to adapt the generated lexicons based on the behavioral theories of trust in face-to-face communications. The generated trustworthiness lexicons are the first lexicons in the context that can be used for future research in the field of online trusting behaviors.

Based on the generated lexicons, online trust reviews were mined to find important characteristics in online trusting decisions from the users' viewpoint. I empirically examined the proposed method by using trust reviews from one of the biggest online social communities, Wikipedia. The results show the superior performance of the proposed method in extracting and classifying trustworthiness characteristics.

The results demonstrate that the developed lexicons based on only the behavioral theories of trust are not sufficient for understanding the trusting behaviors in online social communications. The theory-based trustworthiness characteristics need to be adapted to the domain. The performance of the classifier improved significantly by adapting the theory-based lexicons to the domain.

The proposed method can help in the governance of online communities. Better understanding of user interactions can be gained by mining user opinions about other members. The results of the study can be used in identifying and recommending potential trusted members of the community. Furthermore, the proposed method can be used for studying the trusted communities in the network, trust network structures, and trust propagation.

This study has some limitations that could be addressed in future research. The proposed text mining method is relatively inefficient. The time for lexicon development and some parts of the algorithm increases exponentially with respect to the number of seeds and the review terms. In addition, co-occurrence and Pointwise Mutual Information (PMI) are the only criteria used in classifying trustworthiness characteristics to the related lexicon classes. More criteria, such as semantic similarity techniques, can be considered in the classification algorithm. Furthermore, the importance of the trustworthiness characteristics can be identified and ranked, according to community needs. Finally, future research can study the sentiment of trust reviews and the effect of the sentiment toward each trustworthiness characteristic on the final trust decision.

REFERENCES

Aadithya, K., Ravindran, B., Michalak, T., and Jennings, N. 2010. "Efficient Computation of the Shapley Value for Centrality in Networks," *Internet and Network Economics*, pp. 1-13.

Adamic, L. A., and Adar, E. 2003. "Friends and Neighbors on the Web," *Social Networks* (25:3), pp. 211-230.

Agresti, A. 2013. *Categorical Data Analysis*, John Wiley and Sons, New York.

Ajzen, I. 1985. "From Intentions to Actions: A Theory of Planned Behavior," in *Action Control*. Springer, pp. 11-39.

Ajzen, I. 1991. "The Theory of Planned Behavior," *Organizational Behavior and Human Decision Processes* (50:2), pp. 179-211.

Ajzen, I., and Madden, T. J. 1986. "Prediction of Goal-Directed Behavior-Attitudes, Intentions, and Perceived Behavioral-Control," *Journal of Experimental Social Psychology* (22:5), pp. 453-474.

Al Hasan, M., Chaoji, V., Salem, S., and Zaki, M. 2006. "Link Prediction using Supervised Learning," *SDM06: Workshop on Link Analysis, Counter-Terrorism and Security*.

Anderson, C., and Brion, S. 2010. "Overconfidence and the Attainment of Status in Groups," *Institute for Research on Labor and Employment*.

Anderson, C., Brion, S., Moore, D. A., and Kennedy, J. A. 2012. "A Status-Enhancement Account of Overconfidence," *Journal of Personality and Social Psychology* (103:4), pp. 718-735.

Awad, N. F., and Ragowsky, A. 2008. "Establishing Trust in Electronic Commerce Through Online Word of Mouth: An Examination Across Genders," *Journal of Management Information Systems* (24:4), pp. 101-121.

Ba, S., and Pavlou, P. A. 2002. "Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior," *MIS Quarterly* (26:3), pp. 243-268.

Backstrom, L., and Leskovec, J. 2011. "Supervised Random Walks: Predicting and Recommending Links in Social Networks," *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*: ACM, pp. 635-644.

Bagozzi, R. P., and Yi, Y. 1988. "On the Evaluation of Structural Equation Models," *Journal of the Academy of Marketing Science* (16:1), pp. 74-94.

Baier, A. 1986. "Trust and Antitrust," *Ethics* (96:2), pp. 231-260.

Barber, B. M., and Odean, T. 2001. "Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment," *The Quarterly Journal of Economics* (116:1), pp. 261-292.

Bavelas, A. 1950. "Communication Patterns in Task-Oriented Groups,". *Journal of the Acoustical Society of America* (22:6), pp. 725-730.

Berger, J., Cohen, B. P., and Zelditch Jr, M. 1972. "Status Characteristics and Social Interaction," *American Sociological Review* (37:3), pp. 241-255.

Berger, J., Conner, T. L., and Fisek, M. H. 1982. *Expectation States Theory: A Theoretical Research Program*, Univ Pr of Amer.

Berger, J., Fisek, M. H., Norman, R. Z., Wagner, D. G., Berger, J., and Zelditch, M. 1998. "Formation of Reward Expectations in Status Situations," *Status, Power, and Legitimacy*, pp. 121-154.

Berger, J., Rosenholtz, S. J., and Zelditch Jr, M. 1980. "Status Organizing Processes," *Annual Review of Sociology* (6:1), pp. 479-508.

Bishop, Y. M., Fienberg, S. E., Holland, P. W., Light, R. J., and Mosteller, F. 1977. "Book Review: Discrete Multivariate Analysis: Theory and Practice," *Applied Psychological Measurement* (1:2), pp. 297-306.

Blair-Goldensohn, S., Hannan, K., McDonald, R., Neylon, T., Reis, G. A., and Reynar, J. 2008. "Building a Sentiment Summarizer for Local Service Reviews," *WWW Workshop on NLP in the Information Explosion Era* (14), pp. 339-348.

Blau, P. M. 1964. *Exchange and Power in Social Life*, Transaction Publishers.

Blei, D. M., Ng, A. Y., and Jordan, M. I. 2003. "Latent Dirichlet Allocation, "*Journal of Machine Learning Research* (3), pp. 993-1022.

Borgatti, S. P., and Foster, P. C. 2003. "The Network Paradigm in Organizational Research: A Review and Typology," *Journal of Management* (29:6), pp. 991-1013.

Brandes, U. 2001. "A Faster Algorithm for Betweenness Centrality,". *Journal of Mathematical Sociology* (25:2), pp. 163-177.

Brewer, M. B., and Silver, M. 1978. "Ingroup Bias as a Function of Task Characteristics," *European Journal of Social Psychology* (8:3), pp. 393-400.

Brown, J., Broderick, A. J., and Lee, N. 2007. "Word of Mouth Communication within Online Communities: Conceptualing the Online Social Network," *Journal of Interactive Marketing* (21:3), pp. 2-20.

Buskens, V. 1998. "The Social Structure of Trust," *Social Networks* (20:3), pp. 265-289.

Butler, J. K. 1991. "Toward Understanding and Measuring Conditions of Trust - Evolution of a Conditions of Trust Inventory," *Journal of Management* (17:3), pp. 643-663.

Campos-Castillo, C., and Ewoodzie, K. 2014. "Relational Trustworthiness: How Status Affects Intra-Organizational Inequality in Job Autonomy," *Social Science Research* (44), pp. 60-74.

Carenini, G., Ng, R. T., and Zwart, E. 2005. "Extracting Knowledge from Evaluative Text," *Proceedings of the 3rd International Conference on Knowledge Capture*, pp. 11-18.

Chau, M., and Xu, J. 2012. "Business Intelligence in Blogs: Understanding Consumer Interactions and Communities," *MIS Quarterly* (36:4), pp. 1189-1216.

Chen, G. M., Kim, K. A., Nofsinger, J. R., and Rui, O. M. 2007. "Trading Performance, Disposition Effect, Overconfidence, Representativeness Bias, and Experience of Emerging Market Investors," *Journal of Behavioral Decision Making* (20:4), pp. 425-451.

Chiang, K. Y., Hsieh, C. J., Natarajan, N., Dhillon, I. S., and Tewari, A. 2014. "Prediction and Clustering in Signed Networks: a Local to Global Perspective,". *Journal of Machine Learning Research* (15:1), pp.1177-1213.

Chiang, K. Y., Whang, J. J., and Dhillon, I. S. 2012. "Scalable Clustering of Signed Networks Using Balance Normalized Cut," *Proceedings of the 21st ACM International Conference on Information and Knowledge Management: ACM*, pp. 615-624.

Chiang, K.-Y., Natarajan, N., Tewari, A., and Dhillon, I. S. 2011. "Exploiting Longer Cycles for Link Prediction in Signed Networks," *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*: ACM, pp. 1157-1162.

Chidambaram, L., and Tung, L. L. 2005. "Is Out of Sight, Out of Mind? An Empirical Study of Social Loafing in Technology-Supported Groups," *Information Systems Research* (16:2), pp. 149-168.

Church, K., Gale, W., Hanks, P., and Kindle, D. 1991. "Using Statistics in Lexical Analysis," *Lexical Acquisition: Exploiting Online Resources to Build a Lexicon*, p. 115.

Cohen, J. 1960. "A Coefficient of Agreement for Nominal Scales," *Educational and Psychological Measurement* (1), pp. 37-46.

Colquitt, J. A., Scott, B. A., and LePine, J. A. 2007. "Trust, Trustworthiness, and Trust Propensity: A Meta-Analytic Test of Their Unique Relationships with Risk Taking and Job Performance," *Journal of Applied Psychology* (92:4), pp. 909-927.

Cook, K. S., Levi, M., and Hardin, R. 2009. *Whom Can We Trust?: How Groups, Networks, and Institutions Make Trust Possible*, Russell Sage Foundation.

Dave, K., Lawrence, S., and Pennock, D. M. 2003. "Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews," *Proceedings of the 12th International Conference on World Wide Web*, pp. 519-52.

Ding, X., Liu, B., and Yu, P. S. 2008. "A Holistic Lexicon-Based Approach to Opinion Mining," *Proceedings of the 2008 International Conference on Web Search and Data Mining*, pp. 231-240.

Doney, P. M., and Cannon, J. P. 1997. "An Examination of the Nature of Trust in Buyer-Seller Relationships," *the Journal of Marketing*, pp. 35-51.

DuBois, T., Golbeck, J., and Srinivasan, A. 2011. "Predicting Trust and Distrust in Social Networks," *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom)*: IEEE, pp. 418-424.

Erev, I., Wallsten, T. S., and Budescu, D. V. 1994. "Simultaneous Over- and Underconfidence: The Role of Error in Judgment Processes," *Psychological Review* (101:3), p. 519.

Esuli, A., and Sebastiani, F. 2005. "Determining the Semantic Orientation of Terms Through Gloss Classification," *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*, pp. 617-624.

Everard, A., and Galletta, D. F. 2005. "How Presentation Flaws Affect Perceived Site Quality, Trust, and Intention to Purchase from an Online Store," *Journal of Management Information Systems* (22:3), pp. 55-95.

Fang, X., Hu, P. J.-H., Li, Z., and Tsai, W. 2013. "Predicting Adoption Probabilities in Social Networks," *Information Systems Research* (24:1), pp. 128-145.

Fang, Y. L., Qureshi, I., Sun, H. S., McCole, P., Ramsey, E., and Lim, K. H. 2014. "Trust, Satisfaction, and Online Repurchase Intention: The Moderating Role of Perceived Effectiveness of E-Commerce Institutional Mechanisms," *MIS Quarterly* (38:2), p. 407.

Faraj, S., Jarvenpaa, S. L., and Majchrzak, A. 2011. "Knowledge Collaboration in Online Communities," *Organization Science* (22:5), pp. 1224-1239.

Fiedler, K. 1991. "The Tricky Nature of Skewed Frequency Tables - an Information Loss Account of Distinctiveness-Based Illusory Correlations," *Journal of Personality and Social Psychology* (60:1), pp. 24-36.

Fishbein, M. 1979. "A Theory of Reasoned Action: Some Applications and Implications," *Nebraska Symposium on Motivation*, 27, pp. 65-116.

Fishbein, M., and Ajzen, I. 1975. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*, Reading, Addison-Wesley, MA.

Fiske, S. T., and Taylor, S. E. 1984. *Social Cognition*, Reading, Addison-Wesley, MA.

Foddy, M., Platow, M. J., and Yamagishi, T. 2009. "Group-Based Trust in Strangers the Role of Stereotypes and Expectations," *Psychological Science* (20:4), pp. 419-422.

Fornell, C., and Larcker, D. F. 1981. "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research* (18:1), pp. 39-50.

Freeman, L. C. 1978. "Centrality in Social Networks Conceptual Clarification," *Social Networks* (1:3), pp. 215-239.

Fukuyama, F. 1995. *Trust: The Social Virtues and the Creation of Prosperity*, Free Press Paperbacks.

Funder, D. C. 1987. "Errors and Mistakes - Evaluating the Accuracy of Social Judgment," *Psychological Bulletin* (101:1), pp. 75-90.

Gabarro, J. J. 1978. "The Development of Trust, Influence, and Expectations," *Interpersonal Behavior: Communication and Understanding in Relationships* (290), p. 303.

Gangal, V., Narwekar, A., Ravindran, B., and Narayanam, R. 2016. "Trust and Distrust Across Coalitions: Shapley Value Based Centrality Measures for Signed Networks (Student Abstract Version)," *AAAI*, pp. 4212-4219.

Garton, L., Haythornthwaite, C., and Wellman, B. 1997. "Studying Online Social Networks," *Journal of Computer-Mediated Communication* (3:1), pp. 0-0.

Gefen, D. 2000. "E-Commerce: The Role of Familiarity and Trust," *Omega-International Journal of Management Science* (28:6), pp. 725-737.

Gefen, D. 2002. "Nurturing Clients' Trust to Encourage Engagement Success During the Customization of ERP Systems," *Omega-International Journal of Management Science* (30:4), pp. 287-299.

Gefen, D., and Straub, D. W. 2003. "Managing User Trust in B2C E-Services," *E-Service Journal* (2:2), pp. 7-24.

Gefen, D., Karahanna, E., and Straub, D. W. 2003. "Trust and TAM in Online Shopping: An Integrated Model," *MIS Quarterly* (27:1), pp. 51-90.

Ghani, R., Probst, K., Liu, Y., Krema, M., and Fano, A. 2006. "Text Mining for Product Attribute Extraction," *ACM SIGKDD Explorations Newsletter* (8:1), pp. 41-48.

Giles, J. 2005. "Internet Encyclopaedias Go Head to Head," *Nature* (438:15), pp. 900-901.

Golbeck, J. 2009. "Trust and Nuanced Profile Similarity in Online Social Networks," *ACM Transactions on the Web* (3:4).

Golder, S. A., and Yardi, S. 2010. "Structural Predictors of Tie Formation in Twitter: Transitivity and Mutuality," *Social Computing (SocialCom), 2010 IEEE Second International Conference on*: IEEE, pp. 88-95.

Griffin, D., and Tversky, A. 1992. "The Weighing of Evidence and the Determinants of Confidence," *Cognitive Psychology* (24:3), pp. 411-435.

Guha, R., Kumar, R., Raghavan, P., and Tomkins, A. 2004. "Propagation of Trust and Distrust," *Proceedings of the 13th International Conference on World Wide Web*: ACM, pp. 403-412.

Hair Jr, J. F., Hult, G. T. M., Ringle, C., and Sarstedt, M. 2016. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*, Sage Publications.

Harris, L. C., and Goode, M. M. 2010. "Online Servicescapes, Trust, and Purchase Intentions," *Journal of Services Marketing* (24:3), pp. 230-243.

Harvey, O. 1953. "An Experimental Approach to the Study of Status Relations in Informal Groups," *American Sociological Review* (18:4), pp. 357-367.

Hatzivassiloglou, V., and McKeown, K. R. 1997. "Predicting the Semantic Orientation of Adjectives," *Proceedings of the Eighth Conference on European Chapter of the Association for Computational Linguistics*, pp. 174-181.

Hilbert, M. 2012. "Toward a Synthesis of Cognitive Biases: How Noisy Information Processing Can Bias Human Decision Making," *Psychological Bulletin* (138:2), p. 211.

Hoffman, D. L., Novak, T. P., and Peralta, M. 1999. "Building Consumer Trust Online," *Communications of the ACM* (42:4), pp. 80-85.

Hopcroft, J., Lou, T., and Tang, J. 2011. "Who Will Follow You Back?: Reciprocal Relationship Prediction," *Proceedings of the 20th ACM International Conference on Information and Knowledge Management:ACM*, pp. 1137-1146.

Hu, M., and Liu, B. 2004. "Mining and Summarizing Customer Reviews," *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 168-177.

Hu, M., and Liu, B. 2006. "Opinion Extraction and Summarization on the Web," *AAAI* (7), pp. 1621-1624.

Hulland, J. 1999. "Use of Partial Least Squares (PLS) in Strategic Management Research: A Review of Four Recent Studies," *Strategic Management Journal* (20:2), pp. 195-204.

Javari, A., and Jalili, M. 2014. "Cluster-based Collaborative Filtering for Sign Prediction in Social Networks with Positive and Negative Links," *ACM Transactions on Intelligent Systems and Technology* (5:2), p.24.

Jin, W., Ho, H. H., and Srihari, R. K. 2009. "OpinionMiner: a Novel Machine Learning System for Web Opinion Mining and Extraction," *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1195-1204.

Kane, G. C. 2009. "It's a Network, Not an Encyclopedia: A Social Network Perspective on Wikipedia Collaboration," *Academy of Management Proceedings*, pp. 1-6.

Kane, G. C., Alavi, M., Labianca, G. J., and Borgatti, S. P. 2014. "What's Different About Social Media Networks? A Framework and Research Agenda," *MIS Quarterly* (38:1), pp. 274-304.

Kaplan, A. M., and Haenlein, M. 2010. "Users of the World, Unite: The Challenges and Opportunities of Social Media," *Business Horizons* (53:1), pp. 59-68.

Katz, L. 1953. "A New Status Index Derived from Sociometric Analysis," *Psychometrika* (18:1), pp. 39-43.

Kim, D., and Benbasat, I. 2006. "The Effects of Trust-Assuring Arguments on Consumer Trust in Internet Stores: Application of Toulmin's Model of Argumentation," *Information Systems Research* (17:3), pp. 286-300.

Kim, D., and Benbasat, I. 2009. "Trust-Assuring Arguments in B2C E-Commerce: Impact of Content, Source, and Price on Trust," *Journal of Management Information Systems* (26:3), pp. 175-206.

Kim, S. M., and Hovy, E. 2004. "Determining the Sentiment of Opinions," *Proceedings of the 20th International Conference on Computational Linguistics*, p. 1367.

Kittur, A., and Kraut, R. E. 2010. "Beyond Wikipedia: Coordination and Conflict in Online Production Groups," *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, pp. 215-224.

Kittur, A., Suh, B., Pendleton, B. A., and Chi, E. H. 2007. "He Says, She Says: Conflict and Coordination in Wikipedia," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 453-462.

Kleinberg, J. M., Kumar, R., Raghavan, P., Rajagopalan, S., and Tomkins, A. S. 1999. "The Web As a Graph: Measurements, Models, and Methods," *International Computing and Combinatorics Conference*: Springer, pp. 1-17.

Koh, J., Kim, Y. G., Butler, B., and Bock, G. W. 2007. "Encouraging Participation in Virtual Communities," *Communications of the ACM* (50:2), pp. 68-73.

Komiak, S. Y., and Benbasat, I. 2006. "The Effects of Personalization and Familiarity on Trust and Adoption of Recommendation Agents," *MIS Quarterly* (30:4), pp. 941-960.

Kovelamudi, S., Ramalingam, S., Sood, A., and Varma, V. 2011. "Domain Independent Model for Product Attribute Extraction from User Reviews using Wikipedia," *IJCNLP*, pp. 1408-1412.

Krackhardt, D., Nohria, N., and Eccles, B. 2003. "The Strength of Strong Ties," *Networks in the Knowledge Economy*, p. 82.

Kramer, R. M. 1994. "The Sinister Attribution Error - Paranoid Cognition and Collective Distrust in Organizations," *Motivation and Emotion* (18:2), pp. 199-230.

Kramer, R. M., and Isen, A. M. 1994. "Trust and Distrust: Its Psychological and Social Dimensions," *Motivation and Emotion* (18:2), pp. 105-107.

Kruger, J. 1999. "Lake Wobegon Be Gone! The" Below-Average Effect" and the Egocentric Nature of Comparative Ability Judgments," *Journal of Personality and Social Psychology* (77:2), pp. 221.

Kruglanski, A. W., and Ajzen, I. 1983. "Bias and Error in Human Judgment," *European Journal of Social Psychology* (13:1), pp. 1-44.

Kumar, N., Scheer, L. K., and Steenkamp, J.-B. E. 1995. "The Effects of Supplier Fairness on Vulnerable Resellers," *Journal of Marketing Research*, pp. 54-65.

Kunegis, J., Lommatzsch, A., and Bauckhage, C. 2009. "The Slashdot Zoo: Mining a Social Network with Negative Edges," *Proceedings of the 18th International Conference on World Wide Web*: ACM, pp. 741-750.

Landis, J. R., and Koch, G. G. 1977. "An Application of Hierarchical Kappa-type Statistics in the Assessment of Majority Agreement among Multiple Observers," *Biometrics*, pp. 363-374.

Langer, E. J. 1975. "The Illusion of Control," *Journal of Personality and Social Psychology* (32:2), p. 311.

Leskovec, J., Huttenlocher, D., and Kleinberg, J. 2010a. "Predicting Positive and Negative Links in Online Social Networks," *Proceedings of the 19th International Conference on World Wide Web*: ACM, pp. 641-650.

Leskovec, J., Huttenlocher, D., and Kleinberg, J. 2010b. "Signed Networks in Social Media," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: ACM, pp. 1361-1370.

Levin, D. Z., and Cross, R. 2004. "The Strength of Weak Ties You Can Trust: The Mediating Role of Trust in Effective Knowledge Transfer," *Management Science* (50:11), pp. 1477-1490.

Liben-Nowell, D., and Kleinberg, J. 2007. "The Link-Prediction Problem for Social Networks," *Journal of the American Society for Information Science and Technology* (58:7), pp. 1019-1031.

Lichtenwalter, R. N., Lussier, J. T., and Chawla, N. V. 2010. "New Perspectives and Methods in Link Prediction," *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*: ACM, pp. 243-252.

Lim, K. H., Sia, C. L., Lee, M. K., and Benbasat, I. 2006. "Do I Trust You Online, and If So, Will I Buy? An Empirical Study of Two Trust-Building Strategies," *Journal of Management Information Systems* (23:2), pp. 233-266.

Liu, B. 2012. "Sentiment Analysis and Opinion Mining," *Synthesis Lectures on Human Language Technologies* (5:1), pp. 1-167.

Liu, B. Q., and Goodhue, D. L. 2012. "Two Worlds of Trust for Potential E-Commerce Users: Humans as Cognitive Misers," *Information Systems Research* (23:4), pp. 1246-1262.

Liu, H., Lim, E.-P., Lauw, H. W., Le, M.-T., Sun, A., Srivastava, J., and Kim, Y. 2008. "Predicting Trusts Among Users of Online Communities: An Epinions Case Study," *Proceedings of the 9th ACM Conference on Electronic Commerce*: ACM, pp. 310-319.

Lou, T., Tang, J., Hopcroft, J., Fang, Z., and Ding, X. 2013. "Learning to Predict Reciprocity and Triadic Closure in Social Networks,". *ACM Transactions on Knowledge Discovery from Data (TKDD)* (7:2), p.5.

Lount, R. B., and Pettit, N. C. 2012. "The Social Context of Trust: The Role of Status," *Organizational Behavior and Human Decision Processes* (117:1), pp. 15-23.

Lowry, P. B., Vance, A., Moody, G., Beckman, B., and Read, A. 2008. "Explaining and Predicting the Impact of Branding Alliances and Web Site Quality on Initial Consumer Trust of E-Commerce Web Sites," *Journal of Management Information Systems* (24:4), pp. 199-224.

Ludwig, S., and Nafziger, J. 2011. "Beliefs About Overconfidence," *Theory and Decision* (70:4), pp. 475-500.

Ma, M., and Agarwal, R. 2007. "Through a Glass Darkly: Information Technology Design, Identity Verification, and Knowledge Contribution in Online Communities," *Information Systems Research* (18:1), pp. 42-67.

Madden, T. J., Ellen, P. S., and Ajzen, I. 1992. "A Comparison of the Theory of Planned Behavior and the Theory of Reasoned Action," *Personality and Social Psychology Bulletin* (18:1), pp. 3-9.

Majchrzak, A. 2009. "Comment: Where Is the Theory in Wikis?," *MIS Quarterly* (33:1), pp. 18-20.

Maloney-Krichmar, D., and Preece, J. 2002. "The Meaning of an Online Health Community in the Lives of its Members: Roles, Relationships and Group Dynamics," *Technology and Society* (ISTAS'02), pp. 20-27.

Matsuo, Y., and Yamamoto, H. 2009. "Community Gravity: Measuring Bidirectional Effects by Trust and Rating on Online Social Networks," *Proceedings of the 18th International Conference on World Wide Web*: ACM, pp. 751-760.

Mayer, R. C., and Davis, J. H. 1999. "The Effect of the Performance Appraisal System on Trust for Management: A Field Quasi-Experiment," *Journal of Applied Psychology* (84:1), pp. 123-136.

Mayer, R. C., Davis, J. H., and Schoorman, F. D. 1995. "An Integrative Model of Organizational Trust," *Academy of Management Review* (20:3), pp. 709-734.

McKnight, D. H., and Chervany, N. L. 2001. "Trust and Distrust Definitions: One Bite at a Time," *Trust in Cyber-Societies*. Springer, pp. 27-54.

McKnight, D. H., Choudhury, V., and Kacmar, C. 2002. "Developing and Validating Trust Measures for E-Commerce: An Integrative Typology," *Information Systems Research* (13:3), pp. 334-359.

Mcknight, D. H., Cummings, L. L., and Chervany, N. L. 1998. "Initial Trust Formation in New Organizational Relationships," *Academy of Management Review* (23:3), pp. 473-490.

McKnight, D. H., Kacmar, C. J., and Choudhury, V. 2004. "Shifting Factors and the Ineffectiveness of Third Party Assurance Seals: A Two-Stage Model of Initial Trust in a Web Business," *Electronic Markets* (14:3), pp. 252-266.

McPherson, M., Smith-Lovin, L., and Cook, J. M. 2001. "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology* (27:1), pp. 415-444.

Menon, A. K., and Elkan, C. 2011. "Link Prediction via Matrix Factorization,". *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 437-452.

Metzger, M. J. 2004. "Privacy, Trust, and Disclosure: Exploring Barriers to Electronic Commerce," *Journal of Computer-Mediated Communication* (9:4).

Meyerson, D., Weick, K. E., and Kramer, R. M. 1996. "Swift Trust and Temporary Groups," *Trust in Organizations: Frontiers of Theory and Research* (166), p. 195.

Miller, G. A. 1995. "WordNet: A Lexical Database for English," *Communications of the ACM* (38:11), pp. 39-41.

Mishra, A., and Bhattacharya, A. 2011. "Finding the Bias and Prestige of Nodes in Networks Based on Trust Scores," *Proceedings of the 20th International Conference on World Wide Web*: ACM, pp. 567-576.

Moorman, C., Zaltman, G., and Deshpande, R. 1992. "Relationships Between Providers and Users of Market Research: The Dynamics of Trust Within and Between Organizations," *Journal of Marketing Research* (29:3), p. 314.

Newman, M. E. 2001. "The Structure of Scientific Collaboration Networks," *Proceedings of the National Academy of Sciences* (98:2), pp. 404-409.

O'Madadhain, J., Hutchins, J., and Smyth, P. 2005. "Prediction and Ranking Algorithms for Event-Based Network Data,". *ACM SIGKDD Explorations Newsletter* (7:2), pp. 23-30.

Ordway, E. B. 2009. *Synonyms and Antonyms: An Alphabetical List of Words in Common Use, Grouped with Others of Similar and Opposite Meaning*, Sully and Kleinteich.

Page, L., Brin, S., Motwani, R., and Winograd, T. 1999. "The PageRank Citation Ranking: Bringing Order to the Web," *Stanford InfoLab*.

Peng, W., and Park, D. H. 2011. "Generate Adjective Sentiment Dictionary for Social Media Sentiment Analysis using Constrained Nonnegative Matrix Factorization," *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM-2011).*

Popescu, A. M., and Etzioni, O. 2007. "Extracting Product Features and Opinions from Reviews," *Natural Language Processing and Text Mining*, pp. 9-28.

Preece, J. 2000. *Online Communities: Designing Usability and Supporting Sociability*, Chichester: John Wiley and Sons.

Ransbotham, S., and Kane, G. C. 2011. "Membership Turnover and Collaboration Success in Online Communities: Explaining Rises and Falls from Grace in Wikipedia," *MIS Quarterly* (35:3), pp. 613-627.

Ransbotham, S., Kane, G. C., and Lurie, N. H. 2012. "Network Characteristics and the Value of Collaborative User-Generated Content," *Marketing Science* (31:3), pp. 387-405.

Rempel, J. K., Holmes, J. G., and Zanna, M. P. 1985. "Trust in Close Relationships," *Journal of Personality and Social Psychology* (49:1), p. 95.

Ridgeway, C. L. 2001. "Gender, Status, and Leadership," *Journal of Social Issues* (57:4), pp. 637-655.

Ridgeway, C. L., and Berger, J. 1986. "Expectations, Legitimation, and Dominance Behavior in Task Groups," *American Sociological Review*, pp. 603-617.

Ridgeway, C. L., and Walker, H. A. 1995. "Status Structures," *Sociological Perspectives on Social Psychology*, pp. 281-310.

Ridgeway, C. L., Boyle, E. H., Kuipers, K. J., and Robinson, D. T. 1998. "How Do Status Beliefs Develop? The Role of Resources and Interactional Experience," *American Sociological Review*, pp. 331-350.

Ritschard, G., and Zighed, D. A. 2003. "Goodness-of-Fit Measures for Induction Trees," *International Symposium on Methodologies for Intelligent Systems*, pp. 57-64.

Scaffidi, C., Bierhoff, K., Chang, E., Felker, M., Ng, H., and Jin, C. 2007. "Red Opal: Product-Feature Scoring from Reviews," *Proceedings of the 8th ACM Conference on Electronic Commerce*, pp. 182-191.

Sheppard, B. H., and Sherman, D. M. 1998. "The Grammars of Trust: A Model and General Implications," *Academy of Management Review* (23:3), pp. 422-437.

Sherchan, W., Nepal, S., and Paris, C. 2013. "A Survey of Trust in Social Networks," *ACM Computing Surveys (CSUR)* (45:4), p. 47.

Sia, C. L., Lim, K. H., Leung, K., Lee, M. K., Huang, W. W., and Benbasat, I. 2009. "Web Strategies to Promote Internet Shopping: Is Cultural-Customization Needed?" *MIS Quarterly* (33:3), pp. 491-512.

Song, X., Chi, Y., Hino, K., and Tseng, B. 2007. "Identifying Opinion Leaders in the Blogosphere," *Proceedings of the Sixteenth ACM Conference on Information and Knowledge Management*: ACM, pp. 971-974.

Sproull, L., and Kiesler, S. 1986. "Reducing Social Context Cues: Electronic Mail in Organizational Communication," *Management Science* (32:11), pp. 1492-1512.

Stolte, J. F. 1978. "Power Structure and Personal Competence," *The Journal of Social Psychology* (106:1), pp. 83-92.

Sutton, S. 1998. "Predicting and Explaining Intentions and Behavior: How Well Are We Doing?," *Journal of applied Social Psychology* (28:15), pp. 1317-1338.

Symeonidis, P., Tiakas, E., and Manolopoulos, Y. 2010. "Transitive Node Similarity for Link Prediction in Social Networks with Positive and Negative Links," *Proceedings of the Fourth ACM Conference on Recommender Systems*: ACM, pp. 183-190.

Tang, J., Gao, H., Hu, X., and Liu, H. 2013. "Exploiting Homophily Effect for Trust Prediction," *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining*: ACM, pp. 53-62.

Tang, J., Hu, X., and Liu, H. 2014. "Is Distrust the Negation of Trust? : the Value of Distrust in Social Media," *Proceedings of the 25th ACM Conference on Hypertext and Social Media: ACM*, pp. 148-157.

Tanis, M., and Postmes, T. 2005. "A Social Identity Approach to Trust: Interpersonal Perception, Group Membership and Trusting Behaviour," *European Journal of Social Psychology* (35:3), pp. 413-424.

Thye, S. R. 2000. "A Status Value Theory of Power in Exchange Relations," *American Sociological Review*, pp. 407-432.

Titov, I., and McDonald, R. 2008. "Modeling Online Reviews with Multi-Grain Topic Models," *Proceedings of the 17th International Conference on World Wide Web*, pp. 111-120.

Tsai, W., and Ghoshal, S. 1998. "Social Capital and Value Creation: The Role of Intrafirm Networks," *Academy of Management Journal* (41:4), pp. 464-476.

Turney, P. D. 2002. "Thumbs Up or Thumbs Down? : Semantic Orientation Applied to Unsupervised Classification of Reviews," *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pp. 417-424.

Tversky, A., and Kahneman, D. 1975. "Judgment Under Uncertainty: Heuristics and Biases," *Utility, Probability, and Human Decision Making*. Springer, pp. 141-162.

Varlamis, I., Eirinaki, M., and Louta, M. 2010. "A Study on Social Network Metrics and Their Application in Trust Networks," *Advances in Social Networks Analysis and Mining (ASONAM), 2010 International Conference on*: *IEEE*, pp. 168-175.

Vu, H. T., Le, B. H., and Ho, T.T. 2013. "A Decision-Making Based Feature for Link Prediction in Signed Social Networks," *IEEE International Conference on Computing and Communication Technologies, Research, Innovation, and Vision for the Future (RIVF),* pp. 169-174.

Wagner, D. G., and Berger, J. 1993. "Status Characteristics Theory: The Growth of a Program," *Theoretical Research Programs: Studies in the Growth of Theory* (23), p. 63.

Wang, C., Satuluri, V., and Parthasarathy, S. 2007. "Local Probabilistic Models for Link Prediction," *Seventh IEEE International Conference on Data Mining (ICDM): IEEE*. pp. 322-331.

Wang, D., Pedreschi, D., Song, C., Giannotti, F., and Barabasi, A.-L. 2011. "Human Mobility, Social Ties, and Link Prediction," *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*: ACM, pp. 1100-1108.

Wang, W., and Benbasat, I. 2008. "Attributions of Trust in Decision Support Technologies: A Study of Recommendation Agents for E-Commerce," *Journal of Management Information Systems* (24:4), pp. 249-273.

Wasserman, S., and Faust, K. 1994. *Social Network Analysis: Methods and Applications*, Cambridge University Press.

Watts, D. J., and Strogatz, S. H. 1998. "Collective Dynamics of ‘Small-World’ Networks," *Nature* (393:6684), pp. 440-442.

Webb, T. L., and Sheeran, P. 2006. "Does Changing Behavioral Intentions Engender Behavior Change? A Meta-Analysis of the Experimental Evidence," *Psychological Bulletin* (132:2), p. 249.

Wellman, B. 1997. "An Electronic Group Is Virtually a Social Network," *Culture of the Internet* (4), pp. 179-205.

West, R., Paskov, H. S., Leskovec, J., and Potts, C. 2014. "Exploiting Social Network Structure for Person-to-Person Sentiment Analysis," *Transactions of the Association for Computational Linguistics (TACL)*, PP. 297–310.

Williams, M. 2001. "In Whom We Trust: Group Membership as an Affective Context for Trust Development," *Academy of Management Review* (26:3), pp. 377-396.

Wolf, J. R., and Muhanna, W. A. 2011. "Feedback Mechanisms, Judgment Bias, and Trust Formation in Online Auctions," *Decision Sciences* (42:1), pp. 43-68.

Xiang, R., Neville, J., and Rogati, M. 2010. "Modeling Relationship Strength in Online Social Networks," *Proceedings of the 19th International Conference on World Wide Web*: ACM, pp. 981-990.

Zhang, L., Liu, B., Lim, S. H., and O'Brien-Strain, E. 2010. "Extracting and Ranking Product Features in Opinion Documents," *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, pp. 1462-1470.

Zhang, T., Jiang, H., Bao, Z., and Zhang, Y. 2013. "Characterization and Edge Sign Prediction in Signed Networks," *Journal of Industrial and Intelligent Information* (1:1).

Zheng, X., Wang, Y., Orgun, M. A., Zhong, Y., and Liu, G. 2014. "Trust Prediction with Propagation and Similarity Regularization," *AAAI*, pp. 237-244.

Ziegler, C. N., and Golbeck, J. 2007. "Investigating Interactions of Trust and Interest Similarity," *Decision Support Systems* (43:2), pp. 460-475.

Ziegler, C. N., and Lausen, G. 2005. "Propagation Models for Trust and Distrust in Social Networks," *Information Systems Frontiers* (7:4-5), pp. 337-358.

Zolfaghar, K., and Aghaie, A. 2012. "A Syntactical Approach for Interpersonal Trust Prediction in Social Web Applications: Combining Contextual and Structural Data," *Knowledge-Based Systems* (26), pp. 93-102.

CURRICULUM VITAE

Gelareh Towhidi

Education

Ph.D. Management Information Systems, Minor: Marketing 2018
Sheldon B. Lubar School of Business, University of Wisconsin – Milwaukee
Dissertation Title: Three Essays on Trust Mining in Online Social Networks

M.Sc. Information Technology, Minor: e-Commerce 2006
IUST (Iran University of Science and Technology), Tehran, IRAN
Thesis: A Framework for Selecting the Best e-Business Models for Iranian SMEs

B.Sc. Industrial Engineering, Minor: Planning and System Analysis 2002
Alzahra University, Tehran, IRAN
Thesis: Effective Mechanisms for Bridging the Industry-Academia Gap

Research Summary

Refereed Conference Proceedings

- Towhidi, G. and Srite, M. 2016. "The Effect of Website Quality on Repurchase Intention: The Moderating Role of Cultural Differences," in Proceedings of 22nd Americas Conference on Information Systems (AMCIS 2016), San Diego, CA.
- Towhidi, G. and Sinha, A.P. 2015. "Predicting Opinion Leaders in WOM Communities," in Proceedings of 21st Americas Conference on Information Systems (AMCIS 2015), Puerto Rico.

Conference Presentations

- Towhidi, G. and Zhao, H. 2015. "Are They in Your "Web of Trust" or "Block List"? Trust Prediction in Social Networks," 25th Workshop on Information Technologies and Systems (WITS 2015), Forth Worth, TX.
- Walsh, K J., Thomas, K J., and Towhidi, G. 2013. "The Impact of CEO Compensation Composition and Environment on Merger Exposure," Midwest Academy of Management 56th Annual Conference, Chicago, IL.
- Towhidi, G. and Fathian, M. 2007. "Critical Success Factors for Decision Making on e-Business Model Selection," 5th International Conference in Industrial Engineering, Tehran, Iran.
- Towhidi, G. and Fathian, M. 2006. "A Framework for Selecting the Best e-Business Model for Iranian SMEs," 4th International Management Conference, Tehran, Iran.

Working Papers

- Towhidi, G., Sinha, A.P., Srite, M.D., and Zhao, H. "Trust Formation in Social Networks: A Network-Based Socio-Psychological Model of Online Trusting Behaviors".
- Towhidi, G., Sinha, A.P., and Zhao, H. "Trust Prediction in Social Networks: The Role of Network Structural and Behavioral Trust-Inducing Factors".
- Towhidi, G., Sinha, A.P., and Zhao, H. "Lexicon-Based Trust Mining in Social Networks: Extracting Trustworthiness Characteristics from User-Generated Content".
- Towhidi, G. and Srite, M. "Consumers Online Purchasing Behaviors: The Role of Cultural Differences".
- Towhidi, G. and Ghose, S. "The Economic Impacts of Social Media: Predicting Movie Box-Office Success with Twitter".

Research Interests

- Online Social Networks
- Big Data Analytics
- Data/ Text/ Web Mining
- IS Behavioral Analytics
- Economic Analysis of IS
- E-Commerce

Teaching Experience

University of Wisconsin – Milwaukee, Lubar School of Business

Course Instructor/Adjunct Faculty

BusAdm 537- Enterprise Systems Concepts and Issues	Fall 2017
BusAdm 230- Intro to Information Technology Management	Summer 2016
BusAdm 230- Intro to Information Technology Management	Winter 2016
BusAdm 530- Intro to e-Business	Fall 2015

Computer Lab Leader

BusAdm 536- Business Intelligence	Fall 2016- Spring 2018
BusAdm 230- Intro to Information Technology Management	Spring 2015
BusAdm 230- Intro to Information Technology Management	Fall 2014
BusAdm 230- Intro to Information Technology Management	Summers 2014 - 2015

Teaching Assistant

BusAdm 335- Visual System Development	Fall 2014 - Spring 2015
BusAdm 230 (nline) - Intro to Information Technology Management	Fall 2014 - Spring 2015

Teaching Interests

- Enterprise Systems Concepts and Analytics, SAP
- Business Intelligence
- Predictive Analytics, Data/Text/Web Mining
- Systems Analysis and Design
- Database Development and Management
- e-Commerce Models and Strategies

Academic Activities

University of Wisconsin – Milwaukee
Center for Technology Innovation and SAP University Competence Center

- Project Assistant 2015-Current
- Research Assistant 2013- 2014

Journal/Conference PC Member and Reviewer

- The DATA BASE for Advances in Information Systems
- 44th Academy of Marketing Science Annual Conference, AMS 2016
- Americas Conference on Information Systems, AMCIS 2015 & 2016 & 2018
- International Conference on Information Systems, ICIS 2014

Professional Affiliation

- Association for Information Systems (AIS)
- SAP University Alliance

Training

SAP-ERP Certificate
Graduate Certificate in Enterprise Resource Planning (ERP) / SAP TERP10 certification

Tools

- Programming Languages & Database Systems: Python, Java, VB, Oracle, MySQL Workbench
- Data Analytics Tools: R, SAS, SPSS, Minitab, SmartPLS
- Data/Text/Web Mining Tools: Weka, R, Open NLP, SQL Server Data Mining
- Social Network Analytics Tools: R, Gephi, Pajek, NodeXL, Twitter developer API
- Business Analytics & SAP Tools: Business Warehouse, Business Objects, Crystal Dashboard, Lumira, HANA, Predictive Analytics, Business Explorer, Design Studio