Computational Trust at Various Granularities in Social Networks

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Atanu Roy

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosphy

Jaideep Srivastava
Jisu Huh (co-advisor)

December, 2015

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Prof. Jaideep Srivastava for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a more suitable adviser and mentor for my Ph.D study.

My sincere thanks also goes to my co-advisor Prof. Jisu Huh, who provided me an opportunity to work with her, and who gave access to valuable social science resources. Without her precious support it would not be possible to conduct this research.

Besides my advisors, I would like to thank the rest of my thesis committee: Prof. Loren Terveen, and Prof. Daniel Boley, not only for their insightful comments and encouragement, but also for the hard question which motivated me to widen my research from various perspectives.

I would like to thank all my co-authors Zoheb, Ayush, Muhammad and Brian for agreeing to work with me. Without their work, this thesis would have been incomplete.

Outside University of Minnesota, I would like to acknowledge our collaborators at the Virtual Worlds Observatory project. I would like to thank the faculty members Professor Noshir Contractor, Professor Dmitri Williams and Professor Marshall Scott Poole who opened new horizons for me and taught me how to do inter-disciplinary research. My sincere thanks also goes to Dr. Zhen Wen and Dr. Mercan Topkara from IBM T.J. Watson Research Lab who helped to shape my interest and knowledge in the domain of computational social sciences. I also thank Dr. Lei Wu and Mr. Aaron Ling from Ancestry Data Science, who provided me an opportunity to join their team as a Data Scientist, where I am learning interesting things with hands on experience. I would also like to thank my current team members Dr. Jianlong Qi, Dr. Peng Jiang

and Mr. Jeffrey Sukharev for creating a great data science team at Ancestry where discussion of research ideas is highly appreciated.

I thank my fellow ex and current lab mates, Karthik, Komal, Nishith, Jehwan, Ankit and Dhruv for the stimulating discussions, useful feed backs and for all the fun we have had in the last four years. I would also thank all my close friends, especially Vikas, Vivek, Shurik and Maya from University of Minnesota for making this journey fun and light.

I would like to thank my entire family who taught me everything I know. My father (Alok Kumar Roy) has been instrumental in shaping my outlook towards life and my mother (Sumita Roy) has always been a doting presence. I would like to take this opportunity to individually thank every one in my family and show how important they have been in providing me with all the love and care in this world. Thank you ammu(SmritiKana Roy), kakua (Arup Kumar Roy), boromama(Debesh Roy), mimi(Shibani Roy), dadabhai(Amit Roy), pipi(Supriya Roy), meso(Ratan Lal Ray), kakima(Kakali Roy) and mamonis(Soma, Sharmistha Roy).I would also like to thank my sister in law Tuhina Sarkar, brother in law Shoubho Chakrabarty and ma Uma sarkar, who were always supporting me and encouraging me with their best wishes.

Although I do not believe in life after death, but I still believe that my grandparents(Pradyut, Dwijendra, Gita Roy), chordadu(Pankaj Roy), piso(Amiya Roy) and my mama(Somes Roy) would be elated beyond imagination if they had been alive today to share this moment with me.

I would also like to call out my sisters(Sanjana, Shreetama and Srijata) and brother(Subhojit) and hope that this thesis inspires them to achieve something great in their lives.

Last but not the least, I would like to thank my wife Dr. Chandrima Sarkar. She has always been my best friend, harshest critic, and my greatest appreciator. She has been the greatest contributor to this thesis. Almost every page in this thesis either have her critical feedback or her sharp technical insight. Thank you for choosing me Chandrima.

Dedication

To those researchers whose shoulders this research stands on and to those who will be using this as their stepping stone.

Abstract

Trust has been a ubiquitous phenomenon in human lives. The phenomenon of trust has been studied at various granularities over the centuries by various researchers encompassing all disciplines of academia. Historically, it has been witnessed that the primary mode of studying trust has been surveying subjects and documenting the results. But the burgeoning electronic social media have provided us with the unique opportunity of studying trust under a new perspective, which is known as computational trust. Computational trust is defined as the generation of trust between two human actors mediated through computers. This is an active area of research due to the proliferation of various socially rich datasets over the past decade. This includes massively multi-player online games (MMOs), online social networks and various web services, allowing actors to trust each other in an online virtual setting.

The first part of this thesis investigates various aspects affecting dyadic(or interpersonal) trust, i.e., trust between two actors. This includes formation, reciprocation and revocation of trust. Taking into account various nuances of dyadic trust, this thesis predicts the occurrence of these three phenomena in the datasets. Instead of looking at these phenomena by itself, this thesis looks at this phenomena in conjunction with social relations for better predictive modeling. One of the major requirements in trust applications is identifying the trustworthy actors in the social networks which will be the subject of investigation for the second part of this dissertation. An important factor in the prediction of trust is an actor's inherent ability to trust others and the perception of the actor in the network. This thesis proposes a pair of complementary measures that can be used to measure trust scores of actors in a social network using involvement of social networks. Based on the proposed measures, an iterative matrix convergence algorithm is developed that calculates the trustingness and the trustworthiness of each actor in the network. Trustingness of an actor is defined as the propensity of an actor to trust his neighbors in the network. Trustworthiness, on the other hand, is defined as the willingness of the network to trust an individual actor. The algorithm runs in $O(k \times |E|)$ time where k denotes the number of iterations and $|E|$ denotes the number of edges in the network. This thesis also shows that the algorithm converges to a finite value very

quickly. Lastly, this thesis introduces the concept of "vulnerable paths" and identifies those paths in a social network. Based on the hypothesis that these vulnerable paths are imperative for influence flow, a new algorithm proposed in this thesis, exploits these paths for better and more targeted viral marketing using trust scores. It is shown that there is an improvement as high as 9% in identifying these paths using the proposed algorithm than state of the art trust scoring algorithms.

This thesis makes the following contributions. It studies the generative mechanisms of trust not in isolation, but in conjunction with the social processes(relations) around trust. Whereas earlier studies were interested in looking at the cross-sectional view of trust, this study investigates the longitudinal view of trust. Instead of looking only at the dynamics of initiation of interpersonal trust, this study looks at the various other dynamics such as reciprocation and revocation of interpersonal trust. This study also exploits the negative feedback property in trust to propose computationally stable pair of global trust measures, which can be used to measure the propensity of actors to trust and be trusted in a network. Finally, this pair of scores is leveraged to be used in various applications such as viral marketing, identification of "vulnerable paths" and inoculation of a network from rumor spread.

Contents

List of Tables

List of Figures

TL:DR

People trust each other in social networks. This thesis finds how and why this relationships forms and who are the most probable to form these ones

Chapter 1

Introduction

Trust has been an ubiquitous phenomena in human lives. From time immemorial human society has been based upon trust and social companionship [2, 3].The phenomena of trust has been studied at various granularity over the centuries by various researchers encompassing all fields of academia. It has been studied in great detail from varied disciplines like philosophy [4], social psychology [5], journalism & mass communications [6] and various areas of scientific research [7] like cognitive sciences [8] and trust mediated by computers [9, 10]. In the recent past, trust mediated by computers, better known as computational trust has been the focus of a lot of studies. Most notable among them are the theses by Marsh [10], Golbeck [9] and Ahmad [11]. In the area of trust propagation, researchers have investigated trust metrics motivated by the spreading activation strategies [12] and attack-resistance [13]. In the field of recommender systems, trust metrics have been shown to decrease the error rate [14] of recommendations while inclusion of trustworthiness of users have been shown to improve the effectiveness of recommendations [15]. In the field of multi-agent systems, several trust and reputation metrics have been proposed [16, 17] - such metrics are usually based on past interactions and belief propagation and aggregation over an agent's neighbors.

Historically it has been witnessed that the primary mode of studying trust has been surveying subjects and documenting the results [18, 19]. But the burgeoning electronic social media has provided us with the unique opportunity of studying trust under a new perspective which is known as computational trust. Computational trust is defined as the generation of trust between two human actors mediated through computers. This

is an active area of research due to the proliferation of various socially rich datasets over the past decade. This include massively multi-player online games (MMOs), online social network and various web services allowing actors to trust each other in an online virtual setting. The online social networks provide the users with a real-world like social atmosphere with the advantage that every move that the actors made can be logged and can be used for scientific inquiry.

1.1 Trust in Social Networks

As discussed earlier in this chapter with the advent of online social networks

1.1.1 Online Social Networks as a testbed for Studying Human Relations

From the onset of human history, humans have been social animals. They have hunted, eaten, harvested and settled in groups. By today's definition of social networks all these people have formed their own social networks. But collecting data from these networks have always been the greatest challenge. Thus most of the historical study about human interactions have been done through surveys. Finding survey candidates have always been a challenging task. Moreover verifying the answers is even trickier. There is always an issue of social privacy and motivation which creeps into the subjects' answers in these surveys. And the amount of data collected is only in typically in hundreds.

With the advent of online social networks, humans have embraced them with open hands. It has bridged distances and have brought people together. The biggest advantage of these networks from the point of view of social scientific research is the ability to collect all activities performs hundreds or thousands (& in some cases millions) of respondents. All temporal (every click stream) activities, demographical attributes of actors and group memberships can be studied through these networks.

MMOGs and its Role in Social Science Research

Virtual worlds constitute a class of online environments where millions of people can share a persistent virtual space and interact with one another. Given the many degrees of freedom accorded to players because of the richness of this domain a large number

of behaviours which one observes in the real world are also observed in these virtual environments [20, 21]; these behaviours can be both positive and negative with sufficient similarity to their real counterparts. **M**assively **M**ultiplayer **O**nline **G**ames (MMOs) are a rich class of online games which are analogous to structured virtual worlds. Like the real world, MMOGs are also used for not only gaming purposes, but also for social purposes. As with online social network, these are controlled environments where every action of each users can be archived and can later be studied in great details for social science research. Moreover the engagement offered by these environments are much higher than [22] than the online social networks and the actors are expected to act much more "natural". These environments also provide the garb of anonymity which makes these perfect cauldron to experience all kinds of human social emotions that one may expect in a real world. These games are played by hundreds of thousands of concurrent users and all attributes(demographics) and actions of each user is archived by the system.

The data from one such game called EverquestII was made available by Sony Online Entertainment. The Sony EverQuest (EQ) II game provides an online environment where a vast number of players can log in and coordinate with each other to achieve a particular missions. Note that players are free to invent, choose their mission and to self-organize among groups of their own interest. The game provides several mechanisms such as chat for instantaneously interaction, grouping with several in-game friends to complete quests which are hard to finish single-handedly, trading with fellow players and various others which will be discussed shortly. In this thesis game data set logs is used which was collected over a 35 week period and was completely anonymized. Information from multiple relations were extracted for this thesis and has been presented in the subsequent chapters. The data spans over various servers to make sure all types of activities are captured. Since multiple relations exist between the same individuals, it is known as a multi-relational network. The relations are explained in greater detail in the dataset sections of each chapter of this thesis.

1.2 Representation of Trust & Social Interactions

Abstract social concepts like trust and social interactions are very hard to compute. There are qualitative ways of capturing trust and is achieved through surveys done online or in person [23]. But the pool of surveyees in most of these cases are not large enough to quantitatively deduce patterns or make predictions about formation and revocation of trust [16]. Moreover these surveys are very expensive to conduct both financially and in terms of manual labor required. An alternative is to use proxies of these social phenomena. The underlying assumption is that there exists a "scientific" mapping between the original abstract social concept and the respective proxies chosen [24].

As discussed previously it is of paramount importance for the decision of proxies which represent the original concepts of social interactions and trust. This section discusses the proxies chosen for this thesis to represent trust and social interactions and delves into the reasoning behind these choices.

Proxy for Trust

There are 5 levels of housing access in the EverQuest II. The highest level of access is the *trustee* access where the trustee has almost equal rights as compared to the owner of the house. A trustee can store, touch, move, add, and remove things thus providing with the option of doing anything with the in-game items stored in the house. These items generally take either real money or hours of game time or both for the owner to acquire. Thus the owner's decision of providing trustee access to another player makes him vulnerable to the 2^{nd} person [25], [26]. Thus housing access is used as a proxy for trust. Ahmad in [27] and [28], Borbora in [29, 30], Singhal in [31] and Roy in [32, 22] have previously used the same network as a proxy for trust.

Proxies in other Datasets

Several real life datasets publicly and privately available are used in this thesis. The publicly available datasets are Epinions dataset [33], Slashdot dataset [34], StackOverflow dataset and Twitter retweet dataset. In StackOverflow, an user marking another user's question as "favorite" is considered as a trust link forming between the 2. In

the Twitter "Retweet" network, retweeting refers to the fact that the retweeter trusts the original tweeter's message. Thus, in this dataset, retweeting is used as a proxy for trust. Moreover 2 classical trust datasets from literature, the Epinions and the SlashDot dataset have also been used in this thesis. The details of the these 2 datasets can be found in *Stanford Network Analysis Project*'s dataset collection [1]

1.2.1 Proxy for Social Interactions

Everquest II provides a plethora of in game social activities. These social interactions include grouping, mentoring, chatting and trading. These have been used as proxies for social interactions in this thesis.

1.2.2 Social Patterns & Trust Reciprocation

In this thesis, three broad categories of online virtual social interactions have been investigated namely mentoring, grouping and trading behavior. The primary motive of the inclusion of these networks was to investigate its impact on the formation of trust where housing access is considered to be the proxy for trust.

1.3 Manifestation of Trust in Various Granularities in a Social Network

Trust by definition is dyadic in nature, i.e., between 2 people. Although Ahmad has talked about various other forms of trust in [11], but a closer investigation will yield in the fact that the higher forms of trust are an amalgamation of dyadic trust. Thus when to understand the formative mechanisms of trust, it has to be studied from a dyadic perspective. On the other hand other, properties like scoring of trust in a network and its applications in various domains can only be studied from a global (network-level) perspective. The primary motive of this thesis is to study these aforementioned trust phenomena from various perspectives.

[1] `http://snap.stanford.edu/data/`

1.3.1 Dyadic Trust in a Virtual World

Computation, evolution and prediction of trust in large online social networks are gaining prominence. The importance of trust in any human relationship can be emphasized by the fact that trust between two individuals affect their relationship as a whole. Studying models for formation and reciprocation of trust in a social setting become important not only to understand the propagation of trust in the network, but it also provides us with insights about the nature of social interactions in the network Understanding the dynamics of formation and reciprocation of trust between two people has always been of great interest in the social sciences, including sociology, psychology, and economics. Trust is also fundamental to practically all societal processes, be it commerce, counseling, mentoring, or forming of personal relationships. As our lives move to the digital realm at an ever-increasing pace, understanding the nature of trust becomes even more important, since our-time tested approach of building trust, namely "looking someone in the eye face-to-face" is sometimes being bypassed altogether, e.g. a pair of software engineers working together intensely, but based in diametrically opposite parts of the world, with nary a chance to ever meet in person. This of course has also led to a dramatic increase in confidence games of various sorts to cheat the unaware. Fortunately though, the very same online mechanisms that increase the vulnerability, also provide us an opportunity to study the phenomenon of interpersonal trust at a level of resolution and nuance that was never before possible. A specific example of this is event logs from Massively Multiplayer Online Games (MMOGs), which capture every single event from every player. These events include things that players do (actions) and inter-personal connections they form (relationships). Most MMOGs support a range of actions and relationships, whose goal is to provide players with a "rich real world like" experience. These relationships include an experience where a player may choose to risk "something" that belongs to him to another player. In the MMOG dataset this relationship is investigated is in the form of "housing access" network. This creates risk for the owner, which makes the decision to grant access to a house is a strong marker of trust formation, with the access granter being the "trustor" and the access recipient being the "trustee". This provides a rich dataset for studying various processes that underlie the formation of interpersonal trust between two players, which it is known as "dyadic trust".

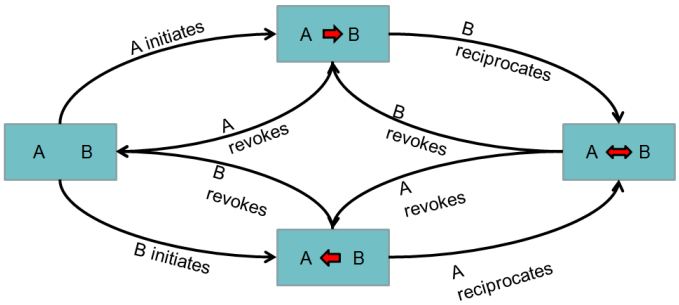


Figure 1.1: State diagramatic view of Dyadic trust

Components of Dyadic Trust

The major components of dyadic trust are formation, reciprocation and revocation of trust. In a trust relation between 2 people, say A & B, when A decides to initiate a trust link towards B (or vice versa), trust is formed between the dyad (2 persons). Initiation of the trust relation is affected by a variety of social interaction factors as will be discussed in a subsequent chapter. When a trust link has formed between a dyad from A to B, and B decides to reciprocate the trust link that A has initiated, reciprocation of trust happens. It may so happen that due to a negative interaction either A or B or both decides to revoke the trust they have accorded to each other. This phenomena is called revocation of trust. A state diagram of these phenomena can be found in figure 1.1.

Formation/Initiation of Trust

Formation of trust in a dyadic setting is a very interesting phenomenon because it is dependent on a plethora of social factors. Since there is a high risk involved in trusting a person, social interactions between the two parties play a major factor indicating whether trust will be formed between the two parties involved. For instance, in the online virtual world, it has been witnessed that formation of trust is generally preceded by a major increase in social interactions which can be conducted with very low risk. These interactions are performed over various networks (like trade, mentor, group) and can only be realized if a multi-relational study of the networks is performed. A study

of only the trust networks fail to capture this rich semantic features and thus end up missing the social nuances of trust formation between 2 individuals.

Recirprocation of Trust

Once trust is accorded to an individual, there are various interesting phenomena which can take place. As discussed previously, one of them is reciprocation. The dynamics of reciprocation varies from network to network depending on the level of barrier for reciprocation. The barrier for reciprocating a trust relationship could be lack of resources or high risk involved. Needless to say, these barriers affect the levels of reciprocation significantly in different networks. For instance, in some networks users have very low barrier level for interacting with each other as there is no commitment from either side to participate in any involved relationship or potential loss. On the other hand, in other networks, the potential for loss is high. It is important to understand questions related to reciprocation across different types of interactions.

Revocation of Trust

A closer study of figure 1.1 will reveal that once trust is formed between 2 persons, several interesting phenomena can take place. One of them reciprocation is discussed in the last paragraph. Another interesting phenomenon that might take place is revocation of trust. Revocation of trust refers to phenomena of taking back trust which has accorded to an individual. This can happen at two stages. After the formation of trust, if trust is revoked, the character dyad (2 persons) returns back to its original state of no trust between. Alternatively when trust is accorded and reciprocation of trust happens, revocation can happen. In one of the cases it might so happen that only a single party revokes the trust, she has accorded and alternatively it may so happen that both the parties revoke their trust in a cascading fashion.

1.3.2 A Network view of Trust

The investigation of dyadic trust culminates into the investigation of global trust in a social network. As already discussed global trust includes methods for scoring trust of each actor in a social network.

Iterative matrix algorithms like HITS [35] is capable of computing pair of scores for each actor in a network. Along with PageRank [36], it tends to prefer nodes with high connectivity. In HITS the two scores positively reinforces each other. As discussed in the aforementioned example, the trust measures should penalize nodes who donate their trust freely. Moreover, the trustworthiness of a person depends on the people who are trusting the actor in question. Unlike HITS and Pagerank whose measurement solely depends on the quantity of links, a mechanism is required where both quality and quantity are considered: The quality of truster along with the quantity of links.

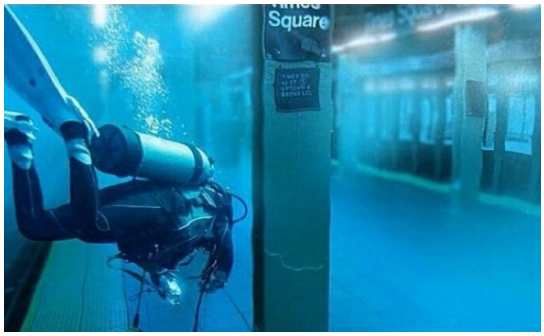
To solve this problem, this thesis introduces two mutually co-related concepts termed as trustingness and trustworthiness. Each actor in the network will be assigned a pair of scores based on the quality and quantity of inlinks and outlinks that the actor has. Two recursive global trust measures are proposed in this paper, namely: trustingness and trustworthiness. Trustingness is defined as the propensity of an individual to trust actors in the network. Trustworthiness is defined as the willingness of the network to trust an actor. As mentioned earlier these measures negatively reinforce each other as displayed in figure 4.1. An actor with high trustingness score "trusts" a lot of actors with low trustworthiness scores. Conversely, an actor with a high trustworthiness score is trusted by a lot of actors with low trustingness scores. The algorithm, **TSM:**, Trust Scores in Social Media, proposed to compute these scores runs in $O(k \times |E|)$ time, where k represents the number of iterations and $|E|$ represents the number of edges in the network. It is shown in the subsequent sections that the result of TSM is bound by a factor $\frac{1}{2^k}$ and the algorithm converges quickly to a stable solution. Using this approach, it becomes very easy to identify nodes with very high trustworthiness (example, trusted news sources like CNN and FOX NEWS in Twitter and Facebook). Moreover, TSM is resistant to actors colluding to increase the trustworthiness score of another actor.

Applications of Trust Scores

Trust Scores as proposed in this thesis can be used in various applications. It was proposed in [37], that trust and influence follows each other closely but in the opposite direction in a network. A trusting B implies that B has some influence over A. Using this hypothesis, trust scores is used to measure influence flow in a network. In this

thesis, a concept called "vulnerable path" based on the trustingness and trustworthiness of neighbors in a network is proposed. A vulnerable path in a social network is loosely defined a path in a social network through which there is a high probability of influence flowing. Using this hypothesis, trust score can be applied in the domain of viral marketing. Identifying the vulnerable paths of influence flow is crucial in the viral marketing. Once these paths are identified marketers can use target marketing to attract the source of these paths and the influence flow inside the network will take care of the diffusion of the product in the network.

Another potential application of trust scores is the ability to inoculate a network from rumor spread. Although this concept is not investigated in this thesis, the author conjectures that the trust scores have the potential of stopping rumor flow in a network. With the advent of online social networks, spreading of rumors through networks has become easier. Figure 1.2 shows a couple popular fake images heavily circulated during Hurricane Sandy. The problem of network inoculation is an anti thesis to the problem viral marketing. Instead of exploiting the vulnerable paths in the network, this problem finds and closes the vulnerable paths in the network. Due to the absence of a suitable dataset, this problem was not investigated in this thesis.



(a) Fake Image of New York City Metro under water

(b) Fake Image of Statue of Liberty washed away by waves

Figure 1.2: Popular fake images circulated in the social media during Hurricane Sandy

1.4 Novelty of this Thesis

This thesis addresses the very important aspects of trust mediated by online networks or better known as computational trust. In the previous studies on computational trust, it has always been studied in isolation. This study has changed the whole landscape and have studied trust along with other networks and impact of them in trust and vice versa. Moreover this thesis has been the first to identify various granularities in computational trust in social networks. This study has also defined the state diagram of dyadic trust (see figure 1.1) and has been instrumental in finding the impacts of other relations on trust formation, reciprocation and revocation.

Moreover this thesis have successfully identified the complementary nature of trust in social networks and have leveraged it by using a simple easy to use iterative matrix convergence algorithm to calculate trust for all actors in a social network.

1.4.1 Thesis Contributions

From the perspective of dyadic trust this thesis makes the following contributions This thesis extends the preliminary work on formation [22] and reciprocation [38] of dyadic trust in an online setting. The contributions in this paper are as follows:

- This thesis provides a complete and nuanced view of dyadic trust in an online virtual setting and build computational models to predict them.
- A detailed prediction model is introduced which improves on the multi-relational features and the aggregation techniques used
- This thesis proposes a new technique to perform time series analysis for finding social patterns preceding and following trust formation.
- A framework capable of modeling computational aspects of trust using established theories from social sciences is developed.
- It is shown that features relating to social interactions are necessary for prediction of trust but are not sufficient.

From the perspective of scoring trust in a social network, this thesis makes the following contributions This research has the following contributions:

- A pair of complementary global trust measures for a social trust network
- A classification system of networks based on risk involved to create links in a network
- Modeling involvement (by a Zipf distribution) and negative feedback property using a decay function
  - Error Bounds of the decay function
- Identify a new technique for a novel viral marketing strategies
- Define the concept of vulnerable edges in a network and use it for applications like viral marketing, network inoculation and weights in influence flow propagation problems.

Overall this thesis makes the following contributions:

- This thesis have identified the social theories and practice used in social science research and have used it to better understand how trust works in humans.
- This thesis furthers the discussion on the information source credibility on the web.
- This thesis makes contributions to the scientific understanding of measuring trust of actors in social netwroks.

Chapter 2

Social Interactions and Trust Formation: A Mutual Reinforcement? An Exploratory Analysis in an Online Virtual Setting

2.1 Overview

Social interactions preceding and succeeding trust formation can be significant indicators of formation of trust in online social networks. This research analyzes the social interaction trends that lead and follow formation of trust in these networks. This enables the author to hypothesize novel theories responsible for explaining formation of trust in online social settings and provide key insights. It is found that a certain level of social interactions threshold needs to be met in order for trust to develop between two individuals. This threshold differs across persons and across networks. Once the trust relation has developed between a pair of characters connected by some social relation (also referred to as a character dyad), trust can be maintained with a lower rate of social interactions.

The first set of experiments is the relationship prediction problem. The emergence of a social relationship like grouping, mentoring and trading between two individuals is predicted over a period of time by investigating the past characteristics of the network. It is found that features related to trust have very little impact on this prediction. In the final set of experiments, the formation of trust between individuals is predicted by looking at the topographical and semantic social interaction features between them. Three semantic dimensions have been generated for this task which can be recomputed with an observed social variable (say grouping) to create a new semantic social variable. In this endeavor, it is successfully shown that, including features related to social interactions, gives an approximate increase of $4 - 9\%$ accuracy for trust relationship predictions.

2.2 Introduction

Computation, evolution and prediction of trust in large online social networks are gaining prominence. The importance of trust in any human relationship can be emphasized by the facts that trust between two individuals affect the relationship as a whole. Studying models for evolution of trust in a social setting becomes important not only to understand the propagation of trust in the network, but it also provides us with insights about the nature of social interactions in the network.

Trust is a ubiquitous phenomenon in human interactions in various social settings and different aspects of trust have been studied across different domains. Golbeck *et. al.* in [39] provides a survey of important research in the field of computational trust and includes models, metrics and applications of social trust. Various models of computational trust have been proposed by [27, 10] and these models seek to formalize the different aspects of trust across domains. In the area of trust propagation, researchers have investigated trust metrics motivated by the spreading activation strategies [12] and attack-resistance [13]. In the field of recommender systems, trust metrics have been shown to decrease the error rate [14] of recommendations while inclusion of trustworthiness of users have been shown to improve the effectiveness of recommendations [15]. In the field of multi-agent systems, several trust and reputation metrics have been proposed [16, 17] - such metrics are usually based on past interactions and belief

propagation and aggregation over an agent's neighbors.

In this study, the relationship between social interactions and trust formation in an online game setting is investigated. Ahmad *et. al.* in [27] has explored various aspects of computational trust in such an online virtual environment. These include specialized and generalized exchange in trust networks [40], relationship of trust to homophily and expertise [28] and trust formation as link prediction [24]. In addition, Borbora *et. al.* in [29] have identified robust predictors of trust in a multi-relational setting. Weekly time-series data for experiments and analysis have been used in this study.

To the best of the author's knowledge, no previous work have investigated the relationship between social interactions and trust in an online multi-relational setting.

This paper makes the following contributions:

- The relationship between social interactions and trust formation in an online virtual setting is investigated, which provides key insights about social patterns required for the formation of trust in these environments. It is discovered that trust formation requires a certain threshold of social interactions. This threshold varies across persons and across networks. But once trust between two parties has been formed, the character pair does not maintain such high levels of social interactions to maintain trust.

- Based on the above insights, both social and trust relationship prediction across multi-relational networks have been set up for this study. To aid in these predictions three semantic dimensions have been proposed to capture the various aspects of social interactions between a character pair. Results from the relationship prediction experiments suggest that social interactions is a good indicator of trust formation but not vice versa.

2.3 Approach

2.3.1 Preliminaries / Assumptions

Abstract social concepts like trust and social interactions are very hard to compute. There are qualitative ways of capturing trust and is achieved through surveys done online or in person [23, 41]. But the pool of surveyees in most of these cases are not

large enough to quantitatively deduce patterns or make predictions about formation and revocation of trust [16]. Moreover these surveys are very expensive to conduct both financially and in terms of manual labor required. An alternative is to use proxies of these social phenomena. The underlying assumption is that there exists a "scientific" mapping between the original abstract social concept and the respective proxies chosen [24].

2.3.2 Problem Statement

Social Patterns

Formally the problem of "finding social patterns preceding and following the creation of trust between two in-game characters" is defined as follows:

Given: A multi-relational social network $G(V, E_1, E_2, \ldots, E_n, E_{trust})$ where each set of edges E_i refers to a social relation in the network. E_{trust} refers to the trust relation in the multi-modal network.

Find out: Social patterns before and after trust is formed between a character dyad in the game.

Assumptions: The proxies of trust and other social interactions have a "scientific" mapping between the abstract concept and the proxies used in the research.

In this research various in-game social networks which are considered proxies for social interactions in the game are provided. In the primary hypothesis described in figure 2.1, an assumption that an increase in social interactions will lead to trust formation which in turn leads to increased social interactions is put forth.

Relationship Prediction

Given: A social network graph $G(V, E)$ where the nodes V represent the actors in the network and edges E represent the existence of a specific relation between them during time t_0 to t_1.

Predict: The existence of a link between two nodes i and $j \epsilon V$ during time interval t_1 to t_2 where $t_2 > t_1 > t_0$.

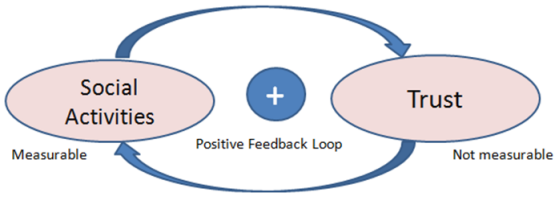


Figure 2.1: social interactions and trust formation: a mutual reinforcement?

2.3.3 Social Patterns & Time Series Clustering

Three broad categories of online virtual social interaction namely mentoring, grouping and trading behaviour were investigated while examining the impact of these relations on the formation of trust. A detailed discussion about the various networks including the trust network will be presented.

First, a time series analysis of the social interaction relation is performed to investigate the social patterns that precedes and succeeds the formation and revocation of trust in these networks. For this investigation, server logs for the social interactions between the characters were collected in an online virtual MMORPG game for a period of a few months.

Next the weekly social interactions history of the players were aligned with the trust relation forming exactly during the halfway through the study interval. This is followed by a time series $k-$means clustering over the data [42, 43].

The plots in figures 2.3(a), 2.3(b) and 2.3(c) show a distinct trend in social interaction before formation of trust in online virtual settings. Although the figures represent impact of separate social interactions on trust formation, the trend of a sharp increase in social interactions immediately before the formation of trust is evident. This leads to hypothesize that a certain threshold of social interactions has to be met before trust can form between two parties. This threshold differs based on individuals & networks.

Once trust is established between the two parties, it requires lesser amount of social interactions compared to formation of trust to maintain the trust between two individuals.

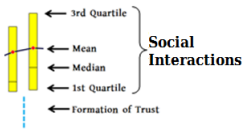


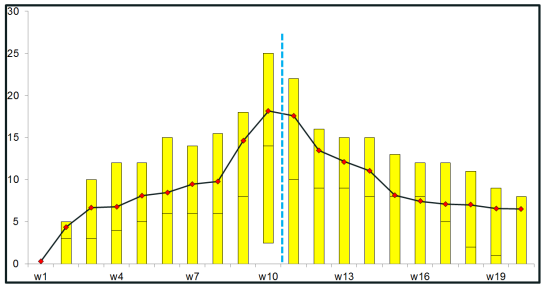
Figure 2.2: Guide for plots in figure 2.3

2.3.4 Relationship Prediction

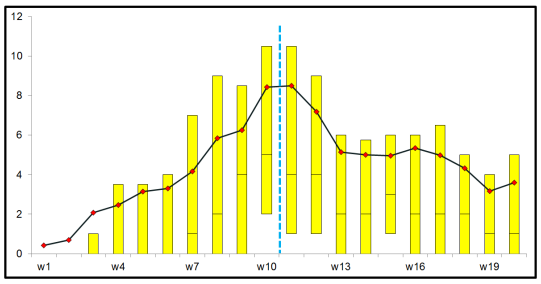
With the insights gained from looking at the social patterns, this study proposes to use these insights for the task of relationship prediction in these networks. In the discussion section, the insights gained from analyzing the social patterns will be reviewed along with the discussion on how it helped in the design of the experiments.

Motivation

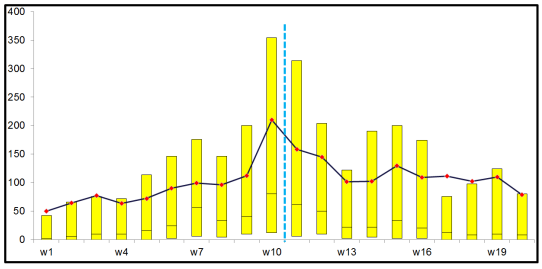
Social patterns and time series clustering provide insights into social interactions before and after trust is formed. The experiments discussed in the subsequent sections provide specific social interaction patterns preceding and following the formation of trust. These trends indicate that formation of trust is accompanied by change in levels of social interactions which affects the hypothesis stated by us in figure 2.1. The hypothesis states that an increase in social interactions will lead to a formation of trust and formation of trust in turn will lead to increased social interactions. Assuming this hypothesis to be true, specific patterns of trust and social interactions will exist during the formation of one another. To test this theory the next task will be to predict the formation of these two relations (social interactions and trust) in a multi-relational setting. To test the effect of social interactions and trust on the formation of each other, this study introduces features pertaining to both in each other's prediction as discussed subsequently. Assuming the primary hypothesis is true, these features should be highly discriminative and should increase the accuracy of the prediction results. Whether they really achieve the feat remains to be seen.



(a) Impact of grouping on trust formation



(b) Impact of mentoring on trust formation



(c) Impact of trading on trust formation

Figure 2.3: The figures refer to the social interaction patterns before and after trust/distrust links are formed between two in-game characters. All interactions are studied over a 20 week period where trust formation between characters form during the 10^{th} week. X-axis refers to the week in question and Y-axis amount of the specific social interaction represented by a box plot. The *blue* dashed vertical line denotes when the trust link was formed between these characters. The index for this figure is shown in figure 2.2

Prediction of Social Relations

In the first prediction task, the effect of trust in prediction of social interactions is investigated. Social interactions, in this research is represented through grouping, mentoring and trading.

Prediction of Trust Relations

In this set of experiments, this study investigates the impact of social interactions on the formation of trust. The control set of the experiment is constructed using the feature sets related to the housing network. To demonstrate the impact of social interactions on trust formation along with the features from the control set, 3 semantic dimensions

[42] are constructed along which the weekly player social interactions is computed to calculate derived semantic features.

2.4 Dataset & Experimental Setup

The Sony EverQuest (EQ) II game provides an online environment where multiple players can log in and coordinate with each other to achieve a particular mission. Note that players are free to invent, choose their mission and to self-organize among groups of their own interest. The game provides several mechanisms such as chat for instantaneously interaction, grouping with several in-game friends to complete quests which are hard to finish single-handedly, trading with fellow players and various others which will be discussed shortly. In this thesis game data set logs is used which was collected over a 35 week period and was completely anonymized. The information needed for the following experiments is extracted from these logs for various interactions. In this section, each of these networks are summarized in terms of the number of nodes and edges, the period of observation, and the direction of edges. The data spans over various servers to make sure all types of activities are captured. Since multiple relations exist between the same individuals, it is referred to as a multi-relational network. The relations are explained below.

1. *Group Network*: There are certain activities and quests in the game which are too difficult for individual players to complete while playing solo. These activities force the players to group together with other players in order to complete these tasks. The resultant network is the grouping network.

2. *Mentor Network:* Mentoring is an in-game feature where more experienced players can *mentor* less experienced players to get them more familiar with the game. The resultant network which connects a mentor to his mentee is known as the mentoring network.

3. *Trade Network:* A trade network is formed by constructing an edge between the two participating entities(players) when they have traded with one another.

it is already mentioned in the assumptions subsection that for quantitative study of trust requires proxies of trust which can be scientifically mapped to the original concept

of trust. In this research, housing access in EverQuestII is identified as a proxy for trust.

1. *House Network:* Every character in the game is entitled to buy in-game houses. Houses serve as a refuge to store in-game virtual items amassed in the game. Moreover it also serves as a place from which a player can sell their goods to other players. Thus from the perspective of in-game wealth, houses are vitally important to their owners. In EverQuest II, a player can *trust* his/her in-game *friend* and allow the person access to his/her house. The friend can view, interact and move objects in and out of these houses. When an owner of a certain house (henceforth referred to as the truster) grants access of his house to an in-game *friend* (henceforth referred to as the trustee), an edge in the housing network is introduced. Granting access to one's house to a different character in the game involves risk since the trustee can "steal" objects from the house which the owner (truster) has put effort to amass. Moreover trusters are allowed to revoke the house accesses from the trustees.

2.4.1 Social Patterns & Time Series Clustering

The server logs from EverQuest II logs all in-game social activities. In this research, the primary aim is to investigate the pattern of activities users *typically* follow before and after granting/revoking trust to a fellow gamer. To accomplish this task, user social interactions data was collected and a time series of the pair's in-game activities for each week in the game was created. Next the time series from all the available character pairs were aligned for a 20 week period. This alignment was done keeping in mind that the formation/revocation of trust between the pair should happen halfway through the 20 week period.

2.4.2 Feature Set Construction

Like any other machine learning technique, feature set selection is important to produce an accurate link prediction model. Features are computed for a pair of nodes i, j.

The experimental feature set can be divided into four broad categories namely topographical [29, 44, 45], homophilic, features related to trust and semantic features [42].

Topographical Features

Topographical features refer to the set of features that exploit the network topology of the underlying network.

Common neighbors This feature identifies the total number of neighbors that are common between any two nodes.

$$\varphi(i,j) = |\Gamma(i) \cap \Gamma(j)| \tag{2.1}$$

Adamic-Adar index Libell-Nowell and Kleinberg in [46] modified the Adamic-Adar index as a feature for link prediction to weigh the neighbours with lower degree more heavily.

$$\vartheta(i,j) = \sum_{k\epsilon(\Gamma(i)\cap\Gamma(j))} \left(\frac{1}{log|\Gamma(k)|}\right) \tag{2.2}$$

Jaccard co-efficient Common neighbor fails to account for the union of the size of the neighborhood of the two nodes. Jaccard's co-efficient considers the union of the size of the neighborhood of the nodes.

$$\zeta(i,j) = \frac{\varphi(i,j)}{|\Gamma(i) \cup \Gamma(j)|} \tag{2.3}$$

Preferential Attachment This is calculated with the premise that a probability of an edge forming between two nodes is proportional to the size of its neighborhood. Preferential attachment is given by

$$\varpi(i,j) = |\Gamma(i)|.|\Gamma(j)| \tag{2.4}$$

Shortest distance Shortest distance calculates the shortest path between any two nodes.

Sum of degree of nodes Sum of degrees adds up the total number of edges incident to both the nodes.

Homophilic Features

Homophilic features are used to describe the properties of nodes in a network.

Sum and Difference of Character Levels MMOGs typically have character level to indicate the in game experience a character has amassed. These features consider the sum and difference of character levels for a given character dyad.

Guild Indicator Guild is an important indicator of homophily.

Trust Feature

This is a binary feature which indicates whether a trust link exists between a character dyad during the period of investigation.

Semantic Features

There is a sharp change in social interactions preceding the formation of trust. In order to capture this sharp change three semantic dimensions are proposed which will be used to recompute weekly player history of an observed social interaction variable, say *number of trade transactions per week*. These dimensions transform the observed social interaction variables to be used during the prediction of trust relationships. In all the 3 semantic dimensions, x_i represents the value of the observed social variable, say *number of trade transactions per week*, for the i^{th} week.

Engagement captures the engagement of a player for the observed variable. For example if engagement is used to recompute the observed variable, say *number of trade transactions per week*, it computes the average number of transactions per week, any two characters made in N number of weeks. Trade engagement for week a is given by.

$$x^a_{engagement} = \frac{1}{N} \sum_{i=a-(N+1)}^{a} x_i \tag{2.5}$$

where x_i represents *number of trade transactions* during the i^{th} *week*.

Intensity captures the ratio of engagement for an observed variable of a node pair compared to their engagement the previous week. In the experiments it is found that

there is a gradual increase in social interactions in the weeks preceding the trust formation. Thus intensity function is weighted to capture this phenomenon by giving the recent weeks more weights.

$$x^a_{intensity} = \sum_{i=a-(N+1)}^{a} i * \left(\frac{x_i}{x_{i-1}}\right) \tag{2.6}$$

A linear weight function is used to generate the results reported in this paper. The function is weighted exponentially based on a modified Katz's co-efficient [47] and was found that linear weighting provided a better accuracy.

Stability This dimension captures the trend of engagement of a player. It has the ability to capture whether there is a decrease or increase in the engagement of a node pair compared to the preceding week. The recent weeks are weighed more heavily using a linear weighting function.

$$x^a_{stable} = \sum_{i=a-(N+1)}^{a} i * Ind\left(x_i, x_{i-1}\right) \tag{2.7}$$

$$Ind\left(x_i, x_{i-1}\right) = \begin{cases} 1 & \text{if } \left(\frac{x_i}{x_{i-1}}\right) > 1, \\ 0 & \text{if } \left(\frac{x_i}{x_{i-1}}\right) = 1, \\ -1 & \text{if } \left(\frac{x_i}{x_{i-1}}\right) < 1, \end{cases}$$

2.4.3 Relationship Prediction

Prediction of Social Interactions

This experiment is designed to test the impact of trust for the prediction of social interactions. The primary set of experiment comprises of the topographical and the homophilic features built on the social interaction networks. This is the control set for the experiment. Next the trust feature is included into the experiment. This experiment is performed thrice; with the trade network, group network and the mentoring network. The data from the 11^{th} week of the year 2006 to the 20^{th} week of the year 2006 is used as the training data and 21^{st} week to 25^{th} week as the test set.

Prediction of Trust Formation

In the social patterns analysis it was observed that a trend emerged which demonstrated that the few weeks preceding the formation of trust there is a certain increase in every type of social interactions between a majorities of character dyads. To exploit this phenomenon, three semantic dimension were defined to capture the social interaction in a sliding window. The semantic dimensions are discussed in the feature set section.

2.5 Experimental Results

2.5.1 Social Patterns & Time Series Clustering

social interactions versus Trust Formation

The three sub-figures of figure 2.3 refers to the average trend of social interaction patterns that leads up to and follows the formation of trust between two in-game characters. The interactions are studied for a 20 week period, 10 weeks leading up to the formation of trust and 10 weeks following the trust has been established. The data is clustered into several behavioral patterns and each of the lines in figure 2.3 refers to the average behavior of the cluster. The average behavior of the population is very similar to the behavior of the largest cluster. In majority of the population it is found that there is an increase in the number of social interactions around the weeks, trust is formed. Few of the clusters exhibit a peak in social interactions before the trust was granted and in a few it is during the week that the trust is granted. For the rest of the clusters, this phenomenon is observed right after trust is granted.

Figures 2.3(a), 2.3(b) and 2.3(c) refer to the average trend of social interaction patterns that leads up to and follows the formation of trust between two in-game characters. The interactions are studied in a sliding window of 20 week period, 10 weeks prior to the formation of trust and 10 weeks following the trust formation. The dotted line in the figure shows the point where trust is formed.

For this analysis, each social interaction network is separately considered and the impact of the interaction on trust formation was investigated. The amount of social interaction for each week was investigated for a period of 20 weeks (sliding window) and is represented in figure 2.3. Each bar in the chart is a part of the “box plot” where

the *y-axis* represents each week's social interaction whereas the *x-axis* represents the week in question. As represented by a box plot, the lower end of each bar is the 1^{st} quartile of social interactions for a week whereas the upper end is the 3^{rd} quartile of the amount the specific social interaction. The blue dashed vertical line is when trust is formed. The average trend is shown as red dots and a trendline is joined across weeks to show the trend of social interactions before and after dyadic trust is formed. This is explained pictorially in figure 2.2.

Each of the group in the sub-figures of figure 2.3 represents the average behavior of a cluster. As discussed in the previous section, for each user, their weekly activity is collected for a period of 20 weeks. A k-means clustering was performed based on the time-series of the weekly user activity over the 20 week period. The expected weekly activity for each cluster is plotted in figure 2.3.

Figures 2.3(a), 2.3(b) and 2.3(c) provide the readers with a visual representation of the social interactions (group, mentor and trade respectively) trends before and after trust formation.

In all the networks, there is always a sharp increase in social interactions for the majority of the population before trust is formed. For a segment of the population a decline in the rate of social interactions can be witnessed immediately after the formation of trust.

Table 2.1: F-measure of various classifiers in predicting different social relations

F-Measure						
	Without Trust			**With Trust**		
	Trade	Group	Mentor	Trade	Group	Mentor
J48	91.32	95.79	95.12	91.25	95.82	94.23
JRip	91.56	96.38	95.32	91.01	95.12	94.79
BayesNet	88.45	89.41	89.01	89.15	89.11	89.55
3-NN	83.69	86.51	85.32	84.1	86.26	85.12

Table 2.2: Table displays the aggregated ranks of all the features used in the task for predicting social interactions with the impact of trust. The rank aggregation algorithm is chosen from [1] where the last column shows the aggregated rank of all the features ranked by the three feature evaluation technique for the relationship prediction across the three networks.

Table: The aggregated ranks of all the features used in for ranking features according to specific statistical properties used in the task for predicting social interactions. The feature having the best value for a particular feature evaluation technique is highlighted.

Network ->		Mentor						Group						Trade						
Attr. Eval Technique ->		Info Gain		Gain Ratio		Chi Square		Info Gain		Gain Ratio		Chi Square		Info Gain		Gain Ratio		Chi Square		Aggregated Rank
Feature Family	Feature	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	
Topo	Common Neighbors	0.3247	3	0.3614	2	3558.43	2	0.4493	2	0.3583	3	3658.23	2	0.3845	2	0.3498	1	3985.32	1	2
	Adamic - Adar	0.425	2	0.3607	3	3558.37	3	0.4125	4	0.3589	2	3325.29	3	0.4125	1	0.2565	4	2748.42	4	3
	Jaccard	0.2264	4	0.3125	4	3452.95	4	0.3847	5	0.3019	5	3012.96	4	0.3568	4	0.2856	3	3189.26	3	4
	Preferential Attachment	0.0812	5	0.2182	5	2431.5	5	0.4325	3	0.3137	4	2958.12	5	0.314	5	0.2123	5	2365.47	5	5
	Shortest Distance	0.5489	1	0.4263	1	4217.23	1	0.5298	1	0.4125	1	4302.1	1	0.3833	3	0.3265	2	3514.23	2	1
	Sum of Degree	0.0583	6	0.1824	6	2086.01	7	0.2201	6	0.2154	6	1352.02	7	0.2458	6	0.1543	7	1874.2	6	6
Homophilic	Sum of Character Level	0.0491	7	0.1231	7	1426.12	9	0.0569	9	0.1865	7	1025.32	8	0.2156	7	0.1865	6	1487.26	7	7
	Difference of Char Levels	0.0212	10	0.0523	10	632.78	10	0.0525	10	0.1802	8	894.56	9	0.1235	9	0.051	10	546.26	8	10
	Guild indicator	0.0222	9	0.1125	8	1823.19	8	0.0895	8	0.051	10	555.49	10	0.1892	8	0.0846	9	889.41	9	9
Trust	Trust Indicator	0.028	8	0.0899	9	2354.21	6	0.185	7	0.1564	9	2125.85	6	0.0889	10	0.0965	8	1025.65	10	8

Table 2.3: This table displays the aggregated ranks of all the features used in the task for predicting trust formation. The original network on which the prediction is performed is the housing network from EverQuest II. The impact networks refer to the impact of specific social relations in the prediction of trust relations. The last column shows the aggregated rank of all the features ranked by the three feature evaluation technique for the relationship prediction across the three impact networks.

Table: The aggregated ranks of all the features used in for ranking features according to specific statistical properties used in the task for predicting trust formation. The feature having the best value for a particular feature evaluation technique is highlighted.

Original Network ->		Housing Trust																	
Impact Network ->		Mentor						Group						Trade					
Attr. Eval Technique ->		Info Gain		Gain Ratio		Chi Square		Info Gain		Gain Ratio		Chi Square		Info Gain		Gain Ratio		Chi Square		Aggregated Rank
Feature Family	Feature	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	
Topo	Common Neighbors	0.5098	3	0.4175	2	2845.61	5	0.4601	4	0.3486	3	3339.96	5	0.4114	4	0.3101	4	3105.25	5	4
	Adamic - Adar	0.4489	6	0.2548	6	1784.13	7	0.316	8	0.1783	8	3100.25	7	0.3483	6	0.2541	5	3001.8	6	6
	Jaccard	0.3012	8	0.2319	7	2841.29	6	0.3749	7	0.2254	7	1845.4	10	0.2789	8	0.1541	7	2741.96	7	7
	Preferential Attachment	0.2746	9	0.1879	8	1325.65	9	0.2147	10	0.1459	9	3335.43	6	0.31	7	0.1011	8	2555.55	8	8
	Shortest Distance	0.5499	1	0.4325	1	3895.12	2	0.4973	2	0.3142	4	4789.12	1	0.4329	3	0.3874	2	4581.21	1	1
	Sum of Degree	0.1545	11	0.0782	10	1012.56	10	0.2674	9	0.0544	12	2749.15	9	0.1823	9	0.0996	9	1401.2	10	10
Homophilic	Sum of Character Level	0.1875	10	0.0556	11	991.45	11	0.1875	11	0.1124	10	1752.32	11	0.0988	11	0.0633	10	1011.1	11	11
	Difference of Char Levels	0.0989	12	0.0412	12	714.83	12	0.16	12	0.0863	11	1000.21	12	0.0663	12	0.0452	12	995.23	12	12
	Guild indicator	0.4193	7	0.1128	9	1548.96	8	0.3985	6	0.2549	6	2841.06	8	0.1474	10	0.0661	11	2101.9	9	8
Semantic Dimensions	Engagement	0.4989	4	0.3415	4	3568.93	3	0.4102	5	0.3785	2	3648.45	4	0.3745	5	0.2113	6	4325.18	2	4
	Intensity	0.5125	2	0.3681	3	4128.21	1	0.5249	1	0.4186	1	4351.01	2	0.4578	2	0.412	1	3589.79	4	1
	Stability	0.4856	5	0.3168	5	3248.69	4	0.4625	3	0.2783	5	4011.96	3	0.4872	1	0.3527	3	4128.32	3	3

Trends & Design Decisions

The plots in figures 2.3(a), 2.3(b) and 2.3(c) shows a distinct trend in social interactions before formation of trust in online virtual settings. Although the figures represent impact of separate social interactions on trust formation, the trend of a sharp increase in social interactions immediately before the formation of trust is evident. This leads the study to hypothesize that a certain threshold of social interactions has to be met before trust can form between two parties. This threshold differs based on individuals & networks.

Once trust is established between the two parties, it requires lesser amount of social interactions compared to formation of trust to maintain the trust between two individuals.

Prediction of Social Interactions

In this section, the information about the multi-relational analysis done in the previous sections is leveraged to study the relational interplay and develop computational models to predict formation and reciprocation of trust.

Motivation

Social patterns and time series clustering provide us with insights into social interactions before and after trust is formed. The experiments discussed in the subsequent sections provide us with specific social interactions patterns preceding and following the formation of trust. These trends indicate that formation of trust is accompanied by change in levels of social interactions which affects the hypothesis stated by us in previous section. The hypothesis states that an increase in social interactions will lead to a formation of trust and formation of trust in turn will lead to increased social interactions. The hypothesis that specific patterns of trust and social interactions will exist during the formation of one another to predict the formation of these two relations (social interactions and trust) in a multi-relational setting is used. Features pertaining to both in each other's prediction as discussed subsequently are introduced in this study.

Feature analysis Results of aggregated ranks of the features involved in prediction of social relations are given in the table 2.2. The feature having the best value for a particular feature evaluation technique is highlighted. Each of the feature evaluation technique produces a ranked listed of attributes. These ranks are aggregated to form the final aggregated ranking using the Borda [48] rank aggregation technique discussed in [1, 49, 50].

A few key insights from the feature analysis are:

- In most of the cases topographical features outperformed the other two families of features.
- It was discovered that the feature indicating the presence of *trust* have an average rank of 8 (out of 10 features).
- Homophilic features performed poorly across all networks and across all feature evaluation techniques.

Prediction Task The results of the the prediction of different social interaction with and without the presence of trust is presented in table 2.1. On a thorough investigation of the results, it becomes evident that inclusion of a trust feature does not have a statistically significant impact on the task of social relationship prediction presented in figure 2.2. In both the prediction tasks represented in tables 2.1 and 2.4, *F1-score* or *F-measure* is used to measure accuracy of the model. Since this study is interested in knowing prediction accuracies of both the positive and negative class F-measure ($F_1 Score$) is used. It is the harmonic mean of *precision* and **recall** score of a classifier is agreed upon as an acceptable measure to calculate the accuracy of a binary classifier [51].

In table 2.1 the results under columns Trade, Group and Mentor signify that the relation in which the link prediction task is performed is Trade, Group and Mentor respectively.

To test the statistical significance of the results, a two sample t-test was performed with an initial null hypothesis stating that "the average accuracy of both the prediction tasks are equal". The p-value for the two sampled t-test resulted in **0.9085** which states that the null hypothesis can not be rejected.

Prediction of Trust Formation

In this section, the results of the impact of social interactions on the prediction of trust formation is presented.

Feature Analysis As discussed earlier, for the prediction of trust relationships, topographical, homophilic and the semantic features were availed . The semantic features are constructed from various observed social variables using semantic dimensions described earlier. On a closer inspection of the detailed results presented in figure 2.3, it is found that the social features defining social impact does very well across all feature evaluation technique and all networks. The topographical features consistently perform well across the board and the homophilic features does poorly.

Feature Set Construction

Like any other machine learning technique, feature set selection is important to produce an accurate link prediction model. Features are computed for a pair of nodes i, j.

Prediction of Trust Formation

In this section, the results of the impact of social interactions on the prediction of trust formation are presented.

Feature Evaluation As with the feature evaluation for social social interaction perdiction, feature evaluation for trust prediction is presented in figure 2.3. For the prediction of trust relationships, topographical, homophilic and semantic features were used. The semantic features are constructed from various observed social variables using semantic dimensions as discussed in previous sections. From figure 2.3, it is ascertained that the social features defining social impact does very well across all feature evaluation techniques and networks. The topographical features consistently outperforms other features and homophilic features do poorly as demonstrated in figure 2.3.

Prediction Task The task of prediction, demonstrated in table 2.4 is divided into 3 distinct tasks. In the first set of prediction task only the topological and homophilic features are used to predict the formation of trust. The results are displayed in table

2.4 under the column name “Without Social Features”. Note that the features for the prediction of trust introduced in section 2.5.1are divided into 3 families: topographical, homophilic and semantic dimensions for social interactions. In the next set of prediction task all the 3 families of features are used and the results are displayed in table 2.4 under the column name “With Topographical + Homophilic + Semantic Dimensions for Social Interactions”. The final prediction task was performed only with the features belonging to the semantic dimensions for social interactions. The results are displayed in the same table under the column name “Only Semantic Dimensions for Social Interactions”. The sub-columns under the last 2 columns, “Trade”, “Group” and “Mentor” indicate the social interaction relation from which the semantic dimensions are created. The last sub-column (T + G) is an aggregated impact of trade and group network on the prediction of trust relations.

Table 2.4: F-measure of various classifiers in predicting trust. The first column represents the control group where trust relations are predicted using topographical and homophilic features. The columns under the "With Topographical + Homophilic + Semantic Dimensions for Social Interactions" column indicate the F-measures of all the 3 families of features. The 3^{rd} column displays the results of the prediction task where only features from semantic dimensions are included. The sub-columns Trade, Group and Mentor indicate the social interaction relation from which the semantic dimensions are created. The last sub-column (T + G) is an aggregated impact of trade and group network on the prediction of trust relations.

F-Measure									
	Without Social Features	**With Topographical + Homophilic + Semantic Dimensions for Social Interactions**				**Only Semantic Dimensions for Social Interactions**			
		Trade	**Group**	**Mentor**	**T + G**	**Trade**	**Group**	**Mentor**	**T + G**
J48	82.26	87.56	89.86	89.32	91.98	77.58	74.68	74.80	79.69
JRip	83.01	88.95	90.12	88.12	92.65	76.23	77.36	74.21	76.99
BayesNet	80.04	84.65	85.65	84.31	86.32	72.75	73.98	72.65	76.08
3-NN	79.65	84.01	84.08	83.21	83.98	70.21	69.54	68.11	72.01

A comparison of the first 2 columns in table 2.4 shows that the F1-score for the second task is much higher that the first task. It can be witnessed that including features related to social interactions heavily influence the results of the prediction task in a positive way. For example, table 2.4 demonstrates the difference in prediction accuracy with and without the inclusion of social interaction features can be as high as 9.64%. The mean F1-score between the first prediction task("Without Social Interactions") was comapred and each sub-task of the second prediction("Trade", "Group", "Mentor", "Trade + Group"). The results are tabulated in the first row of table 2.5. To confirm whether the difference of F-measure score between the prediction tasks are statistically significant, one sided Welch two sample t-test was performed where the alternative hypothesis states that the true difference in average F-measure score is greater than 0. The P-values are tabulated in the 2^{nd} row of table 2.5. It can be seen that out of 4, in 2 cases the p-values are < 0.01 and in all cases they are < 0.02. Thus it can be concluded that including features relating to social interactions lead to a statistically significant improvement of trust prediction.

Next the results of the full feature set were compared to only social features (presented in column 3 of figure 2.4). It can be seen that there is a comparable difference between the 2 results. The difference between the average F-measure scores is 13.29% and the p-value for the alternative hypothesis that "the average F-measure of the full feature classifier is higher than the "only social features set" is $2.339 * 10^{-13}$ provided in table 2.5.

Table 2.5: Statistical Comparison of Trust Prediction tasks

	Trade	**Group**	**Mentor**	**T + G**
Mean Difference	5.05	6.19	5.16	7.49
P-Value	0.00750	0.00896	0.01781	0.01588

2.6 Discussion & Future Work

This brings the discussion back to the initial conundrum. *social interactions and trust formation. Are they mutually dependent on each other?* So does the hypothesis introduced in figure 2.1 holds? This research was started with the intuition that a healthy

positive social interaction between a pair of individuals builds up to trust. This trust in turn leads to more social interactions.

On a closer look at the results it can be found that the first part of the intuition holds. Not only positive social interaction leads to trust, but it is one of the essential features in the formation of trust. Although this might be counter-intuitive, but the preliminary investigation suggests that the feedback loop does not hold. This study demonstrates numerous instances and trends (summarized in figure 2.3) that for a the majority of the population, formation of a trust link is followed immediately by a sharp drop in social interactions. The results suggest that trust is dependent on social interactions whereas the reverse does not hold.

The preliminary hypothesis of trust leading to the strengthening of social interactions in an online virtual setting is refuted by this research. A possible explanation for this sharp decrease can be attributed to the fact that formation of trust is the motivating factor for a high rate of social interactions between the two parties. Once trust is formed between the two individuals the motivation for keeping such a high rate of social interactions diminishes thereby leading to a decrease in the rate of social interactions. This is based on the concept of social bandwidth proposed by Robin Dunbar in [52]. Social bandwidth is defined as the amount of resources a person has for social interactions in their life. Everyone has a limited social bandwidth [52] and before a trust link is formed it can be conjectured that the two level of social interactions increases between the two persons in order to test whether the trustee can be trusted. Once trust is formed between the two this motivation disappears and both of them can invest their social bandwidth with other friends.

To test the hypothesis that social interactions is a strong predictor of trust formation and not the other way around, this study proposes to predict both trust and social interactions relationship and measure the effect of each other in these predictions using binary supervised link prediction techniques. The first set of experiments, whose results are shown in figures 2.2 and 2.3, check the performance of features across EQ II dataset for both prediction tasks. Several attribute evaluation techniques are employed and in both the experiments the topographical features perform exceptionally. In the prediction of social interactions, the feature which is the indicator of trust performs very poorly across all networks and across all attribute evaluation techniques. This

provides the readers with some evidence that trust will not be a “good” predictor of social interactions. In case of social interactions being a predictor of trust, it was found that along with the topographical features, the features related to social interactions performs well across all networks and all attribute evaluation techniques. This is in sharp contrast with the previous result where the feature related to trust performed poorly everywhere. This provides the readers with the intuition that features related to social interactions will considerably affect the prediction of trust in these networks.

The next task involved the prediction of these relationships across the in-game virtual networks. During the prediction of social interactions, as demonstrated in table 2.1, the inclusion of a trust feature does not improve the prediction of social relations in the multi-relational networks. This leads us to conclude that trust is not a good predictor of social interactions. On the other hand, during the prediction of trust relation across different networks, as is evident from table 2.4, it was noticed that the features related to social interactions improves the prediction accuracy of the trust prediction considerably. Across all networks and classifiers there is a significant increase in prediction accuracies, when the features related to social interactions are included. This experiment provides the readers with a hypothesis that features related to social interactions are good predictors of trust but not vice versa. The prediction task confirms this hypothesis. This reinforces the belief in the fact that social interactions clearly impacts the formation of trust whereas vice versa does not hold true.

It is very hard for researchers to obtain data pertaining to trust formation and revocation. Most of the public trust datasets are merely reputation metrics in which the owner(truster) has nothing at stake. For example in the Epinions dataset [14] used for trust prediction, users can express their ”Web of Trust”, i.e. reviewers whose reviews and ratings they have consistently found to be valuable and their ”Block list”, i.e. a list of authors whose reviews they nd consistently offensive, inaccurate, or not valuable. In this dataset the users are rating other users without putting anything at stake. The author feels that this is more of a reputation metric than a trust metric. Moreover the publicly available trust datasets are not multi-relational, i.e., they do not possess social interactions trends. In the future, if these kind of multi-relational datasets are available, the authors theorizes that these results can be generalized to trust formation in other settings.

The question of whether trust leads to social interactions or vice versa can open new vistas of research. Topics like the impact of social interactions on the evolution of trust can be a very interesting and rewarding field to investigate. Does social interactions impact the evolution of trust the same way it does the formation of trust? Or does social interactions impacts evolution of trust in ways very different from what is seen in this research. These are questions which only a thorough investigation of these phenomena can answer.

Chapter 3

Reciprocation and Revocation of Dyadic Trust

3.1 Introduction

The rapid growth in the amount and richness of online interactions, through Massively Multi-player Online Games (MMOGs) such as EverQuest II [1] and World of Warcraft 3 [2] arecreating social interaction data at an unprecedented scale. As mentioned earlier, these virtual worlds provide a rich environment for studying user interactions and have been used in several recent experimental studies [53, 54, 55, 56]. Moreover these datasets provide high resolution and long period data about social interactions and therefore are very useful for detailed empirical analysis. In this chapter, the analysis is scoped towards reciprocation of dyadic trust in EverQuest II (EQ2) MMOG dataset and provide interesting insights about the trust dynamics in the EQII environment. Modeling abstract human concept such as trust is challenging [40] . Although, it will be greatly interesting to develop computation models to study human trust, however this study is restricted to the analysis and inferences of the proxy of trust in the EQ2 dataset due to the following practical challenges related to studying human trust.

1. Trust, an abstract concept, is very hard to model. Modeling trust requires identification of a proxy of trust which has a scientific mapping to the original concept.

[1] https://www.everquest2.com
[2] http://us.battle.net/wow/en/

2. Trust between 2 persons form as a result of the several types of social interactions between them. It is paramount that trust is studied not by itself but along with these social interactions which are factors that influence trust. A primary challenge involved in this research is a longitudinal study of multi-relationships that an entity engages in. In practice it is difficult to gather such personal information involving human subjects.

3. Multi-relational datasets, where one of the relationships is trust is hard to find and in itself is non- trivial. Moreover identification of the proxy for trust along with the identification of the right social interactions to factor in are quite challenging.

4. Trust between 2 persons (dyadic trust) can be broken down into several phenomena. For example formation, reciprocation and revocation of trust. Although they belong to the same abstract concept, the generative mechanism of each of them is very different from the other. The primary challenge of this research lies in correctly identifying factors affecting these phenomena and proposes right approaches to model them.

This chapter reports results from a detailed empirical study on the basic processes underlying the reciprocation and revocation of interpersonal trust between pairs of players in an online game setting, which is also known as "dyadic trust". As already discuused above, the data used is the full player logs from EverQuestII for around 675,000 players, collected over a ten month period, and represents natural gameplay without any intervention, and hence is akin to a natural experiment. This chapter makes a number of contributions. First, it quantifies the connection between degree of social interactions and its impact on trust reciprocation, i.e. the second player in a dyad reciprocates the trust accorded to him. A key finding is that a certain threshold of social interactions is needed before this happens, with the threshold itself dependent on the personality characteristics of the players involved, i.e. how trustworthy or trusting each is. It is observed that reciprocation is not automatic but depends on various factors including degree of social interactions, homophily with the 1st player, and the players social status, e.g. centrality index. A key observation is that (i) in a dyad with the players having significantly different social status, the one with the lower status is the one to first express trust, and (ii) this often causes the higher status player to take notice, which is

observed as more mentoring, trades, group activities etc. being initiated by the higher status player. However, in very few of these cases does this "checking out of the lower status player by the higher status player" result in actual trust being accorded by the higher to the lower. This is the first quantification of the "scaffolding role" played by lower familiarity threshold relationships such as chat and trade, in the formation of high familiarity relationships such as reciprocated trust. The general approach used for testing the ideas explored in this study is to build predictive models, of trust reciprocation and revocation, and then testing them which is possible because of the availability of longitudinal data. With unprecedented amounts of behavioral data becoming available, the approach presented is a very promising way to build more nuanced models and further the community's understanding of dyadic trust.

The second part of this chapter deals with the revocation of trust in a dyadic setting which can be considered as the final state of dyadic trust. Trust revocation is a very hard problem to study for various reasons. The primary reason being availability of datasets. Since revocation of trust has a negative connotation attached to it, it is hard to find datasets that denotes proxies which can be used as proxies for trust revocation. Fortunately the multi-relational dataset from EverQuestII used to study the phenomena of trust formation and reciprocation contains a proxy for revocation for trust. It is the same proxy used in the study of formation and reciprocation of trust which is the relation of providing and revoking housing access to in-game friends. Building models for trust formation and reciprocation primarily depends on the quantity of (social and topographical) interactions that 2 individuals have in the network. Thus using the meta data present in multi relational social network datasets works well while building these models. But the phenomena of revocation is very different. More than quantity, trust revocation depends on the quality of social interactions. One bad experience on the part of an actor can lead to trust revocation which in itself is very hard to capture from a dataset that contains only the metadata about the social interactions.

Understanding the dynamics of reciprocation and revocation of trust between two people has always been of great interest in the social sciences, including sociology, psychology, and economics. Trust is also fundamental to practically all societal processes, be it commerce, counseling, mentoring, or forming of personal relationships. As our lives move to the digital realm at an ever-increasing pace, understanding the nature of trust

becomes even more important, since our-time tested approach of building trust, namely "looking someone in the eye face-to-face" is sometimes being bypassed altogether, e.g. a pair of software engineers working together intensely, but based in diametrically opposite parts of the world, with nary a chance to ever meet in person. This of course has also led to a dramatic increase in confidence games of various sorts to cheat the unaware. Fortunately though, the very same online mechanisms that increase the vulnerability, also provide us an opportunity to study the phenomenon of interpersonal trust at a level of resolution and nuance that was never before possible. A specific example of this is event logs from Massively Multiplayer Online Games (MMOGs), which capture every single event from every player. These events include things that players do (actions) and inter-personal connections they form (relationships). Most MMOGs support a range of actions and relationships, whose goal is to provide players with a "rich real world like" experience. A unique feature in most MMOGs is the ability of a player to create a house in which personal items, acquired either via many hours of play effort or payment of actual money, can be stored. Further, the owner can decide to give someone else access to their house, sometimes with sufficient privileges to allow the latter to move items out of the house, potentially without informing the owner. This creates risk for the owner, which makes the decision to grant access to a house is a strong marker of trust formation, with the access granter being the "trustor" and the access recipient being the "trustee". This provides a rich dataset for studying various processes that underlie the formation of interpersonal trust between two players, which is known as "dyadic trust".

Once trust is accorded to an individual, there are various interesting phenomena which can take place. The dynamics of reciprocation varies from network to network depending on the level of barrier for reciprocation. The barrier for reciprocating a trust relationship could be lack of resources or high risk involved. Needless to say, these barriers affect the levels of reciprocation significantly in different networks. For instance, in some networks users have very low barrier level for interacting with each other as there is no commitment from either side to participate in any involved relationship or potential loss. On the other hand, in other networks, the potential for loss is high. It is important to understand questions related to reciprocation across different types of interactions.

A closer study of figure 3.1 will reveal that once trust is formed between 2 persons, several interesting phenomena can take place. One of them reciprocation is discussed in the last paragraph. Another interesting phenomenon that might take place is revocation of trust. Revocation of trust refers to phenomena of taking back trust which has accorded to an individual. This can happen at two stages. After the formation of trust, if trust is revoked, the character dyad (2 persons) returns back to its original state of no trust between. Alternatively when trust is accorded and reciprocation of trust happens, revocation can happen. In one of the cases it might so happen that only a single party revokes the trust, she has accorded and alternatively it may so happen that both the parties revoke their trust in a cascading fashion.

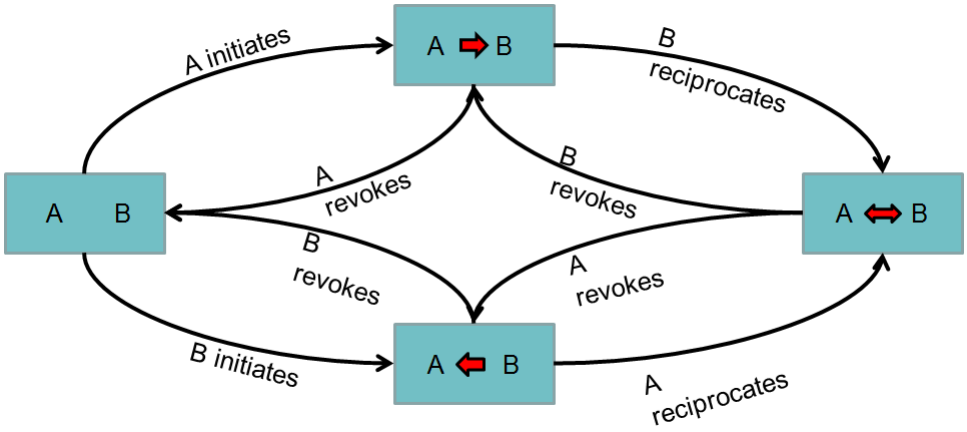


Figure 3.1: State diagramatic view of Dyadic trust

To summarize the dynamics of complex network relationships cannot be studied in isolation because social interactions may play a critical role in building the trust formation or reciprocation.

3.1.1 Contributions

This chapter is a continuation of the preliminary study of dyadic trust [22] in an online setting. The contributions in this chapter are as follows:

- This chapter provides a complete and nuanced view of dyadic trust in an online

virtual setting and build computational models to predict them.

- This chapter introduces a detailed prediction model for trust reciprocation and studies the multi-relational features and the aggregation techniques.
- This chapter builds a framework capable of modeling computational aspects of trust reciprocation and revocation using established theories from social sciences.

3.2 Related Works

The notion of reciprocation as defined by Gouldner [57] is the norm that people should help those who help them.

Researchers have studied reciprocation in great detail [58] in the field of sociology. By definition reciprocation states that people tend to help those who help them. If there tends to be a positive social interaction between the acorded and the acordee the relationship tend to be reciprocated. This has been studied in great detail in the fields of online social networks [59], organizational support [60], and anthropology [61]

As with other sociological study, most studies of reciprocation are performed with a handful of surveyees [62, 63] most of whom are undergraduate students from universities. and used to understand specific human behavioral aspect such as happiness [64] or altruism [63]. However such studies do not focus on trust reciprocation in a dyadic relationship. Although these studies primarily focus on reciprocation of human relationship, they do it primarily on a single relation. They do not use one relationship to predict the reciprocation in a different relationship as in a multi-relationship setting.

3.3 Dataset

As mentioned earlier the dataset used for this chapter originates from the logs of a Massively Multiplayer Online Role Playing Game (MMORPG) called EverQuestII. EverQuestII is an online environment where multiple players can log in and coordinate with each other to achieve a particular mission. Every interaction that a player make with the environment is logged and can be used later to analyze. The game provides various mechanisms to interact with the environment and with other players present in

the immense virtual world. Each interaction is logged and recorded. For example when a group of players team up to complete a mission (say of killing a "unkillable" monster), the underlying relation formed is called group network. The same set of networks is used in the last chapter as proxies for trust and social interactions.

To reiterate the networks used in this chapter are housing access network, which is used as a proxy for trust. The proxies for social interaction as was used in the previous chapter are group network, mentor network and trade network. A detailed discussion of these can be found in the previous chapter.

3.4 Approach

3.4.1 Trust Reciprocation in Housing Access Network

Trust reciprocation of a network is defined as the phenomena where forward trust edges (say player A trusts player B) have a corresponding backward edge (player B reciprocates the trust accorded to A). This is represented in figure 3.1.

As stated in the previous chapter the housing access network is considered as a proxy for trust. Only a total of 14.0% of the forward connections in the housing network is reciprocated back. The average response time for a reciprocation is 27.03 days and the response distribution is shown in figure 3.2. It can be observed that the response time distribution follows a variation of the power law.

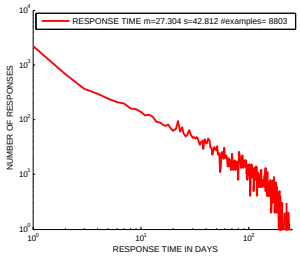


Figure 3.2: Distribution of the times of trust reciprocation in the EverQuest II dataset

3.4.2 Assumptions

Abstract social concepts like trust and social interactions are very hard to compute. There are qualitative ways of capturing trust and is achieved through surveys done online or in person [23]. But the pool of surveys in most of these cases are not large enough to quantitatively deduce patterns or make predictions about formation and reciprocation of trust [16]. Moreover these surveys are very expensive to conduct both financially and in terms of the manual labor required. An alternative is to use proxies of these social phenomena. The underlying assumption is that there exists a "scientific" mapping between the original abstract social concept and the respective proxies chosen [24].

3.4.3 Proxies

As discussed previously it is of paramount importance for the decision of proxies which represent the original concepts of social interactions and trust. This section discusses the proxies that is choosen in this research to represent trust and social interactions and discuss the decisions behind these choices.

Proxy for Trust

One of the previous section discusses the EverQuest II dataset that is for the purpose of this research. Section 2.4 discusses in detail the mechanism of housing access in the game. As previously discussed there are 5 levels of housing access in the game. The highest of access is the *trustee* access where the trustee has almost equal rights as compared to the owner of the house. A trustee can store, touch, move, add, and remove things thus providing with the option of doing anything with the in-game items stored in the house. These items generally take either real money or hours of game time or both for the owner to acquire. Thus the owner's decision of providing trustee access to another player makes him vulnerable to the 2^{nd} person [25], [26]. In this study, the trustee level of housing access is used as a proxy for trust. Ahmad in [27, 28], Roy in [32] and various other authors have previously used the same network as a proxy for trust.

Proxy for Social Interactions

Everquest II provides a plethora of in game social activities as discussed in the dataset section. These social interactions include grouping, mentoring, chatting and trading. These are used as proxies for social interactions in this research.

3.4.4 Social Patterns & Trust Reciprocation

This study has investigated two broad categories of online virtual social interactions namely grouping and trading behavior while examining the impact of these relations on the reciprocation of trust where "trustee" housing access is considered to be the proxy for trust.

First, a time series analysis of the social interaction relation is performed to investigate the social patterns that precedes and succeeds the reciprocation and revocation of trust in these networks.

The whole population is divided into 2 sub-populations, one in which the phenomena of trust reciprocation is observed and the rest in which it is absent. The weekly social interactions history of the 2 sub-population of players are aligned so that for every dyad, trust is initiated after 6 weeks. In other words, the social interaction patterns are studied for a period of 6 weeks before trust initiation for the entire population. Now for the population in which trust has not been reciprocated, the next 16 weeks of social interaction is studied in order to differentiate the behavior when compared to the population where trust is reciprocated. For the second population, where trust is reciprocated, the period between initiation and reciprocation is divided into 4 sub buckets. Since the average period of reciprocation is approximately 4 weeks (figure 3.2), the period is divided into 4 sub divisions. Once trust is reciprocated, the behavior is studied over a period of next 6 weeks.

Next a time series analysis is performed over the data. Time series analysis enables the research in analyzing the data in a longitudinal fashion and can identify the social interaction trends that precede and follow the initiation and reciprocation of trust. This provides the readers with a quantitative way to interpret how social interactions affect the reciprocation of trust in an online virtual setting.

3.5 Patterns of social interactions versus trust reciprocation

First, a time series analysis is performed on the social interaction relation to investigate the social patterns that precedes and succeeds the initiation and reciprocation of trust. The dynamics of trust reciprocation is different from the problem of trust formation. Thus the social patterns analysis was designed differently for the problem of trust reciprocation. The problem of trust reciprocation is defined as the point where the 2^{nd} actor decides to give back trust to the primary actor. Although the problem can be modeled as the problem of trust formation but an extra fact that trust has been accorded can be leveraged in the problem of trust reciprocation. In the case of trust formation(initiation), only the social interactions preceding the trust formation can be modeled into a predictive analysis. In this case the social interactions can be further subdivided into 2 categories. The first category is the set of social interactions that the dyad did before the first trust was accorded in the pair. The previous chapter on trust formation demonstrated that there has been a sharp rise in social interaction between users before trust is accorded. And once trust is accorded the interactions fall off. But in the case where reciprocation really does happen, according to the last chapter there should be another spike of social interactions. Thus the second set of social interactions features for the problem of trust reciprocation is investigation the social interaction patterns once trust has been accorded but has not been reciprocated.

In this chapter, an analysis is performed where the weekly social interactions history of the players are aligned with the formation and reciprocation of trust. The whole of social interactions between 2 characters are viewed within a 16 week period. For every pair of characters that have ever formed a trust relationship in the EverQuest II dataset is divided into two categories. First, the pair of players for whom trust has been accorded but never reciprocated. Second, the group of characters where trust has been reciprocated. It is seen in [38] that the average time required to reciprocate trust in the EverQuestII dataset is 27.304 days which can be approximated to 4 weeks. For the character dyads that reciprocated the trust accorded, the whole interaction between the formation and reciprocation of trust is compressed into buckets which are eventually represented as weeks. For those who have not reciprocated trust, the social interactions

between them is analyzed for the next 10 weeks. The analysis is modeled as a time series analysis. Time series analysis helps in analyzing the data in a longitudinal fashion and can identify the social interaction trends that precede and follow the formation [32] and reciprocation of trust. This provides a quantitative way to interpret how social interactions affect the reciprocation of trust in an online virtual setting.

Trends & Design Decisions

The plots in figures 3.3(a) and 3.3(b) show us a distinct trend in social interaction before formation of trust, between formation and reciprocation of trust and after reciprocation in online virtual settings. Although the figures represent impact of separate social interactions on trust formation, the trend of a sharp increase in social interactions immediately before the formation of trust, an immediate dip and again a spike just before reciprocation is evident. This leads us to hypothesize that sustaining a high amount of social interaction after trust initiation usually leads to reciprocation of trust.

But for those pairs in which reciprocation is not observed, once trust is established, the amount of social interactions is not maintained at such high levels.

3.5.1 The Problem of Predicting Trust Reciprocation

With the insights gained from looking at the social patterns, this study proposes to use these insights for the task of relationship prediction in these networks. In the discussion section, the insights gained from analyzing the social patterns will be reviewed to determine how it helps in designing the experiments for this study.

3.5.2 Problem Statement

Given: A social network graph $G(V, E)$ where the nodes V represent the actors in the network and edges E represent the existence of a specific relation between them during time t_0 to t_1.

Predict: The existence of a link between two nodes i and $j \epsilon V$ during time interval t_1 to t_2 where $t_2 > t_1 > t_0$.

Motivation

Social patterns and time series from figure 3.3 clustering provide us with insights into social interactions before and after trust is reciprocated. The experiments discussed in the subsequent sections provides with specific social interaction patterns preceding and following the reciprocation of trust. The trends as discussed in section 3.5 forwards the hypothesis that "sustaining a high amount of social interaction after trust initiation usually leads to reciprocation of trust". To test this theory the next task will be to predict the reciprocation of these two relations (social interactions and trust) in a multi-relational setting. To test the effect of social interactions on trust reciprocation of each other, features pertaining to both are introduced in the subsequent section. Assuming the primary hypothesis is true, the social interaction features between the formation and reciprocation should be highly predictive in nature. Whether they really achieve the feat remains to be seen.

3.6 Predicting Trust Reciprocation

In this section, a computational model to predict a high barrier relationship, such as a trust, using information about the medium barrier interactions between the nodes (players) is presented. The empirical analysis in the section 3.5 showed that the success (completion) of the trust relationship is influenced by the magnitude of trade/group activities between the two players involved in developing mutual trust relationship. This experiment is further extended in this section to quantitatively evaluate the impact of the of medium barrier interactions such as trade to predict high barrier relationships such as trust reciprocation. To derive any conclusions from this experiment, the analysis is performed for period of 9 months. To make the experiment more realistic in terms of various interactions, several other features (described below) are added. The features are further subdivided into 3 families: namely **Topographical**, **Homophilic** and **Social Semantic**. A detailed discussion of each of the family is provided below:

Topographical Features

Topographical features refer to the set of features that exploit the network topology of the underlying network.

Let us assume $\Gamma(i)$ represents the local neighborhood of a vertex i.

Common neighbors This feature identifies the total number of neighbors that are common between any two nodes.

$$\varphi(i,j) = |\Gamma(i) \cap \Gamma(j)| \tag{3.1}$$

Adamic-Adar index Libell-Nowell and Kleinberg in [46] modified the Adamic-Adar index as a feature for link prediction to weigh the neighbours with lower degree more heavily.

$$\vartheta(i,j) = \sum_{k\epsilon(\Gamma(i)\cap\Gamma(j))} \left(\frac{1}{log|\Gamma(k)|}\right) \tag{3.2}$$

Jaccard co-efficient Common neighbor fails to account for the union of the size of the neighborhood of the two nodes. Jaccard's co-efficient considers the union of the size of the neighborhood of the nodes.

$$\zeta(i,j) = \frac{\varphi(i,j)}{|\Gamma(i) \cup \Gamma(j)|} \tag{3.3}$$

Preferential Attachment This is calculated with the premise that a probability of an edge forming between two nodes is proportional to the size of its neighborhood. Preferential attachment is given by

Shortest distance Shortest distance calculates the shortest path between any two nodes.

Sum of degree of nodes Sum of degrees adds up the total number of edges incident to both the nodes.

Homophilic Features

Homophilic features are used to describe the properties of nodes in a network.

Sum and Difference of Character Levels MMOGs typically have character level to indicate the in game experience a character has amassed. These features consider the sum and difference of character levels for a given character dyad.

Guild Indicator Guild is an important indicator of homophily.

Trust Feature

This is a binary feature which indicates whether a trust link exists between a character dyad during the period of investigation.

Semantic Features

There is a sharp change in social interactions preceding the formation of trust. In order to capture this sharp change, three semantic dimensions are proposed in this study which will be used to recompute weekly player history of an observed social interaction variable, say *number of trade transactions per week*. These dimensions transform the observed social interaction variables to be used during the prediction of trust reciprocation relationships. In all the 3 semantic dimensions, x_i represents the value of the observed social variable, say *number of trade transactions per week*, for the i^{th} week.

Engagement captures the engagement of a player for the observed variable. For example if engagement is used to recompute the observed variable, say *number of trade transactions per week*, it computes the average number of transactions per week, any two characters made in N number of weeks. Trade engagement for week a is given by.

$$x^a_{engagement} = \frac{1}{N} \sum_{i=a-(N+1)}^{a} x_i \tag{3.4}$$

where x_i represents *number of trade transactions* during the i^{th} *week*.

Intensity captures the ratio of engagement for an observed variable of a node pair compared to their engagement the previous week. In the experiments it is found that there is a gradual increase in social interactions in the weeks preceding the trust formation. Thus the weighted intensity function is used to capture this phenomenon by giving the recent weeks more weights.

$$x^a_{intensity} = \sum_{i=a-(N+1)}^{a} i * \left(\frac{x_i}{x_{i-1}} \right) \tag{3.5}$$

A linear weight function is used to generate the results reported in this study. The function is weighted linearly based on a modified Katz's co-efficient [47] since it was found that linear weighting provided a better accuracy.

Stability This dimension captures the trend of engagement of a player. It has the ability to capture whether there is a decrease or increase in the engagement of a node pair compared to the preceding week. The recent weeks are weighed more heavily using a linear weighting function.

$$x^{a}_{stable} = \sum_{i=a-(N+1)}^{a} i * Ind\left(x_i, x_{i-1}\right) \tag{3.6}$$

$$Ind\left(x_i, x_{i-1}\right) = \begin{cases} 1 & \text{if } \left(\frac{x_i}{x_{i-1}}\right) > 1, \\ 0 & \text{if } \left(\frac{x_i}{x_{i-1}}\right) = 1, \\ -1 & \text{if } \left(\frac{x_i}{x_{i-1}}\right) < 1, \end{cases}$$

Use of Semantic Features

The semantic dimensions of *Engagement*, *Intensity* and *Stability* is converted into features by combining them with social interactions relations like grouping, trading and mentoring. Thus each social interaction is converted into 3 social semantic. For example, the social interaction *trade* can converted to $Intensity_{trade}$. For the purpose of reciprocation these features are further subdivided into 2 families. Social semantic features before formation of trust and social semantic features between trust formation and reciprocation.

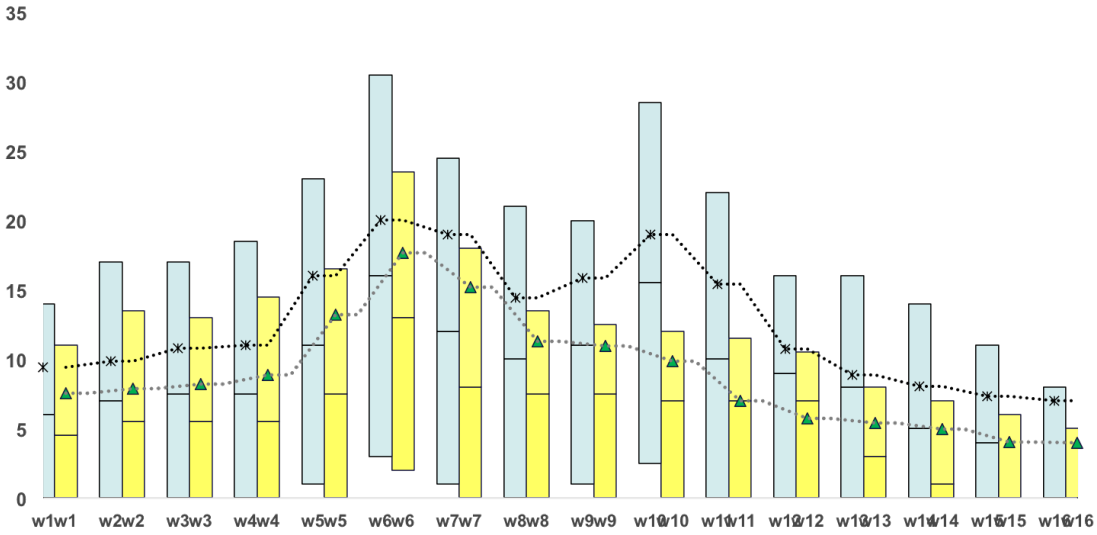
3.6.1 Prediction Model

As mentioned earlier, the aim of this experiment is to quantitatively compare the impact of different features (described above) to predict trust reciprocation between two nodes. In the previous section, it is hypothesized that the success of trust reciprocation (completion) can be determined by the amount of medium barrier interactions between player A and B. This hypothesis is validated in this experiment using a computational model for prediction. The trust reciprocation problem is considered as a binary class

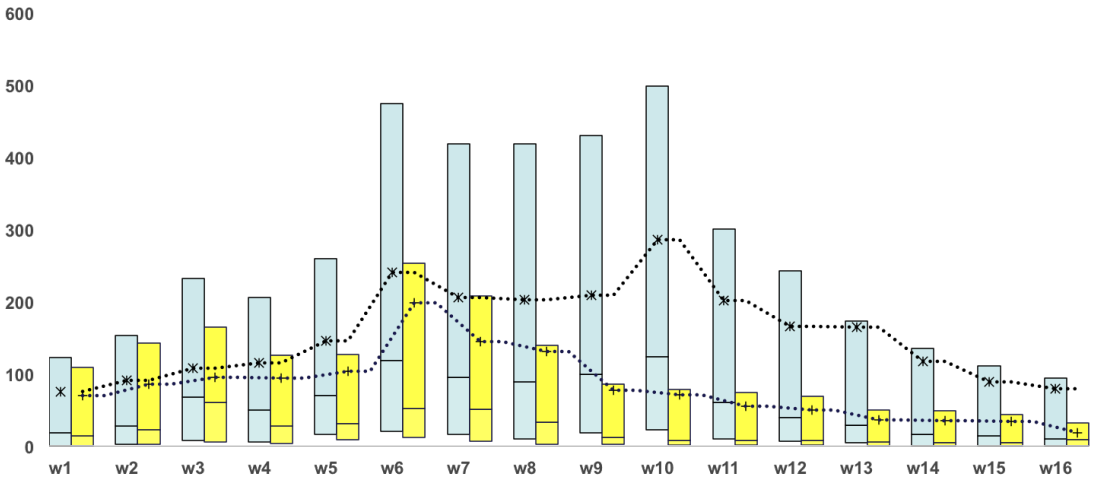
prediction problem. J48 decision tree is used as the binary class prediction model for predicting trust reciprocation between a pair of nodes using the feature sets (discussed above) for that pair of nodes.

The experiment is divided into 3 sub parts like its formation counterpart. The first experiment was performed without the social semantic features. The next set of experiment is performed with the complete set of features namely, the topographical features, the homophilic features, and the 2 families of social semantic features as discussed in the last subsection. To test the primary hypothesis put forth, the last model is created by removing the social semantic feature family before the formation of trust. This helps in proving the effectiveness of social interactions patterns between trust formation and reciprocation and their predictive powers in predicting the phenomena of reciprocation.

3.7 Results & Discussion



(a) Impact of Grouping on trust reciprocation



(b) Impact of Trading on trust reciprocation

Figure 3.3: The figures refer to the social interaction patterns between trust links are formed and are reciprocated. This figure is a comparison of social interaction patterns of dyad that have and have not reciprocated trust. For this study 6 weeks of interactions before trust formation is studied and 6 weeks of interaction after trust reciprocation is studied. The time between formation and reciprocation was divided into 4 buckets and the interactions were divided into those buckets. The index for this figure is shown in figure 3.4

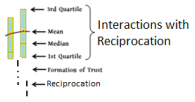


Figure 3.4: Index for figure presented in figure 3.3.

Table 3.1: F-measure of various classifiers in predicting trust. The first column represents the control group where trust is predicted using topographical and homophilic features. The columns under the "Complete Model" column indicate the F-measures of all the 3 families of features. The 3^{rd} column displays the results of the prediction task where social semantic features from trust formation is removed. The sub-columns Trade, Group and Mentor indicate the social interaction relation from which the semantic dimensions are created. The last sub-column (T + G) is an aggregated impact of trade and group network on the prediction of trust relations.

F-Measure									
	Without Social Features	**Complete Model**				**Removing Social Semantic Features before Trust Formation**			
		Trade	**Group**	**Mentor**	**T + G**	**Trade**	**Group**	**Mentor**	**T + G**
J48	0.7932	0.8567	0.8661	0.8511	0.8748	0.8486	0.8745	0.8245	0.8541
JRip	0.8112	0.869	0.8342	0.8812	0.9019	0.8745	0.83	0.8756	0.8979
BayesNet	0.8096	0.8663	0.8456	0.8742	0.8999	0.8845	0.871	0.8743	0.8898
3-NN	0.7661	0.8312	0.8215	0.8777	0.8652	0.8741	0.8122	0.8212	0.8785

A comparison of the 3 classes of models presented in table 3.1 portrays the difference in classification accuracies of the various families of models proposed in a previous section. It can be seen clearly that the introduction of social semantic features have considerably improved the F-1 score of the entire prediction task. The readers would like to compare the results provided in the first column with the second column to do that. Next the hypothesis that the behavior of a dyad after trust is accorded is a good indication can be seen from comparison of 2^{nd} and the 3^{rd} columns of table 3.1. It can be seen that the mean differences between the 2 set of models is very low which leads to the strengthening of the aforementioned hypothesis.

3.8 Trust Revocation

The third part of the "dyadic trust" puzzle is the problem of trust revocation as can be seen in figure 3.1. After the discussion of the two most important phenomena of dyadic trust namely formation and reciprocation, this study investigates the problem of trust revocation to complete the state diagram introduced in figure 3.1. Trust revocation is a phenomenon where one takes away the trust that she had accorded to another individual. Like formation and reciprocation, trust revocation was investigated using the compelling social interaction factors that affect the revocation of trust.

For this analysis, the study separately considers each social interaction network and check the impact of the interaction on trust revocation. A check is performed on the amount of social interaction for each week for a period of 20 weeks and is represented in figure 3.5.

It is very hard to study the behavior of each character pair in the game. Based on empirical analysis, it was decided to cluster the entire character pairs into 3 behavioral clusters to investigate the effect of social interactions on trust formation. The data is clustered into several behavioral patterns and each of the lines in figure 3.5 refers to the average behavior of the cluster. The average behavior of the population is very similar to the behavior of the largest cluster. Since the number of trust revocation pairs are very low compared to the formation and reciprocation problem, a different method is introduced to study the social trends for the revocation problem.

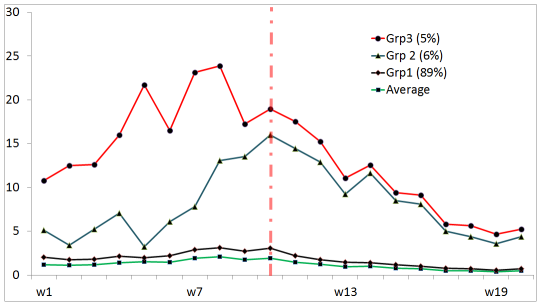
3.8.1 Social Interactions versus Trust Revocation

Figures 3.5(a), 3.5(b) and 3.5(c) provide the readers with visual patterns of the social interactions (group, mentor and trade respectively) trends before and after trust revocation. It has already been discussed that EverQuest II allows its users to revoke trust which provides the unique opportunity to investigate the social patterns before and after revocation of trust. It was found that the social interaction patterns before revocation of trust comparable to that of formation of trust in each network although the magnitude of social interactions were much higher (approximately 5 times) in case of trust formation.

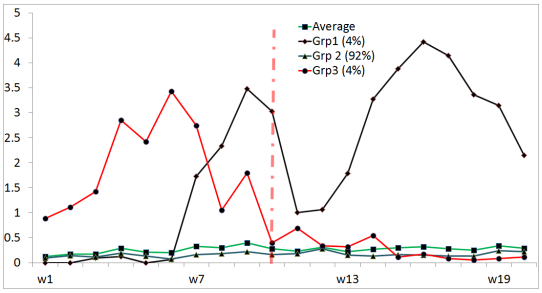
3.9 Experiments & Results for Revocation Study

3.9.1 Experimental Setup

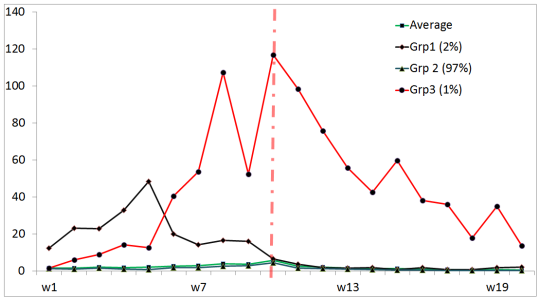
The experiments for the revocation were modeled exactly the same way as in the reciprocation problem. Sections 3.6 and 3.6.1 provide a detailed discussion of the feature set used and the experimental setup. The only difference in these 2 cases are the 2 distinct populations created. In case of reciprocation the 2 populations referred to the population of users who have reciprocation trust and those who have initiated but not reciprocated trust. In this case the 2 population refer to as the population who have revoked trust and those who have initiated but have not revoked trust. Otherwise the features, and the models of prediction all remain the same and can be studied from the previous sections.



(a) Impact of grouping on trust revocation



(b) Impact of mentoring on trust revocation



(c) Impact of trading on trust revocation

Figure 3.5: The figures refer to the social interaction patterns before and after trust revocation between two in-game characters. All interactions are studied over a 20 week period where trust/distrust between characters form during the 10^{th} week. X-axis refers to the week in question and Y-axis refers to the **average** number of social interaction session (as defined in the last section) of each cluster in question. The whole population of in-game characters in the dataset were clustered into 3 behaviorial categories and the colored lines in the plot represents the **average behavior** of a **single behaviorial cluster**. The "Average" describes the mean behavior of the **entire population**. The *red* dashed and dotted vertical line denotes the week where the trust revocation link was formed between these characters. The percentages in the parenthesis next to each group refers to the percentage of the total population that belongs to a certain group.

Table 3.2: F-measure of various classifiers in predicting revocation of trust. The first column represents the control group where trust relations is predicted using topographical and homophilic features. The columns under the "Complete Model" column indicate the F-measures of all the 3 (Topographical, Homophilic and Social Semantic) families of features. The 3^{rd} column displays the results of the prediction task where features from trust formation are excluded. The sub-columns Trade, Group and Mentor indicate the social interaction relation from which the semantic dimensions are created. The last sub-column (T + G) is an aggregated impact of trade and group network on the prediction of trust relations.

F-Measure									
	Without Social Features	**Complete Model**				**Removing Social Interaction Features before Trust Formation**			
		Trade	**Group**	**Mentor**	**T + G**	**Trade**	**Group**	**Mentor**	**T + G**
J48	0.412	0.698	0.694	0.662	0.749	0.512	0.541	0.465	0.543
JRip	0.432	0.662	0.673	0.672	0.716	0.547	0.52	0.452	0.512
BayesNet	0.401	0.645	0.618	0.692	0.691	0.423	0.412	0.412	0.489
3-NN	0.354	0.638	0.605	0.606	0.646	0.552	0.441	0.422	0.474

3.9.2 Results

The results presented in table 3.2 compared to the prediction tasks performed in the trust formation (table 2.4) and reciprocation (table 3.1) is very different. Although the prediction accuracies for the entire model is decent, it is far off from the prediction accuracies that is demonstrated in the last 2 prediction tasks. This leads us to believe that prediction a negative emotion like revocation is tougher and more nuanced than a positive interaction like trust formation or reciprocation. A detailed discussion about this will be provided in the following subsection.

3.9.3 Discussion

Trust revocation is a negative phenomenon. It is hypothesized that it is harder to predict than reciprocation and formation. Phenomena like formation and reciprocation depends on the quantity and quality of interactions between individuals. Thus it is relatively easier to predict from a dataset which necessarily contains metadata and not individual interactions. In this scenario, metadata means that the dataset contains information like who mentored whom and when. But it does not contain the social interaction details that happened during the session. On the other hand revocation is a negative phenomenon. It is hypothesized that revocation solely depends upon the quality of interactions. One terrible experience can easily lead to a revocation which is hard to model using a quantitative, metadata-based dataset used for this research. Moreover revocation is a much rarer phenomenon compared to the two other phenomena. For example, in a 14 week period there has been only 1035 revocation instances compared to 36578 non-revoked instances. The number of data points for revocation is too little to build a comprehensive nuanced model for trust revocation.

Chapter 4

Trustingness & Trustworthiness: A Pair of Complementary Trust Measures in a Social Network

4.1 Overview

The increase in analysis of real life social networks has led to a better understanding of the ways humans socialize in a group. Since trust is an important part of any social interaction, researchers use such networks to understand the nuances of trust relationships. One of the major requirements in trust applications is identifying the trustworthy actors in these networks. This chapter proposes a pair of complementary measures that can be used to measure trust scores of actors in a social network using involvement of social networks. Based on the proposed measures, an iterative matrix convergence algorithm is developed that calculates the trustingness and the trustworthiness of each actor in the network. Trustingness of an actor is defined as the propensity of an actor to trust his neighbors in the network. Trustworthiness, on the other hand, is defined as the willingness of the network to trust an individual actor. The algorithm is proposed based on the idea that a person having higher trustingness score contributes to the trustworthiness of its neighbors to a lower degree. Conversely, a higher trustworthiness score is

a result of lots of neighbors linked to the actor having low trustingness scores. The algorithm runs in $O(k \times |E|)$ time where k denotes the number of iterations and $|E|$ denotes the number of edges in the network. Moreover, the study shows that the algorithm converges to a finite value very quickly. Finally this study uses the proposed scores for trust prediction in various social networks and show that the proposed algorithm performs better (average 5%) than the state of the art trust scoring algorithms.

4.2 Introduction

The previous decade has seen the emergence of social networks representing every sphere of life. There are applications which do not primarily depend on such networks, but build those as a result of actors interacting with each other. Websites like Facebook, Google+ and Twitter are examples of applications where the users interact directly with each other thereby creating a network of their own. These networks are extremely large with Facebook reaching a billion users in the recent past. Actors, in these networks, connect directly with each other, share videos and audio, and perform a host of other engaging activities. There is a second class of online applications where the primary motive of the application is not to directly interact with each other but to use a specific service that the host is offering. For example, sites like Youtube and Dailymotion are popular video hosting sites whereas sites like Epinions, SlashDot and Reddit let users rate products and movies and generate, edit and read content. Although the primary motive is not to interact, these applications incentivize actors to interact by commenting, trusting and/or liking each other. There is a third variety of networks which are formed as a result of the actors playing online computer games. These games are very engaging in nature [65] and are considered a microcosm of the real life society. On one hand these games provide a platform for millions of players to share a concurrent virtual world and interact with the objects and on the other hand it allows the players to interact within themselves.

As discussed previously, the network formed in each of the applications varies vastly. The edges represent various concepts in each of the networks. For example, in Facebook, an edge between two actors can represent the fact that they have befriended each other whereas an edge in Youtube can represent an actor liking or commenting on a

second person's video. There are networks which captures trust also. For example, in Epinions.com, users are allowed to "trust" each other. The interactions of a user and the "web of trust" thus formed, determine the reviews that a user finally sees[1] .

Trust is an abstract human concept. There are various connotations of trust and each form substantially differs from the other. Moreover, there is the element of human perception. Being an abstract concept, trust cannot be measured directly. The only way to measure trust is to identify proxies which can be scientifically mapped to the abstract concept of trust [22]. The strength of trust depends on both the edge weight between the "truster" and the "trustee" and also the inherent propensity of the "truster" to trust actors in the network. For example, let us assume, in a network of several actors, there is one actor who trusts almost every other actor in the network. Whereas a second person in the network is persnickety about the actor he chooses to trust. Thus, the trust conferred by the second person will be more valuable to an average actor in the network.

As can be seen, datasets having edges representing human relations is very common in social media. These datasets provide a very rich medium to study human relations like trust. It becomes highly important for various disciplines to identify the actors in these networks who are "highly trustworthy" and those who trusts a lot of fellow actors. Identifying and scoring these actors not only help in classical problems like trust prediction in social networks but also help in solving problems like "stopping rumor spread" and "viral marketing". This chapter discusses in detail how scoring trust in a social network aids in trust prediction.

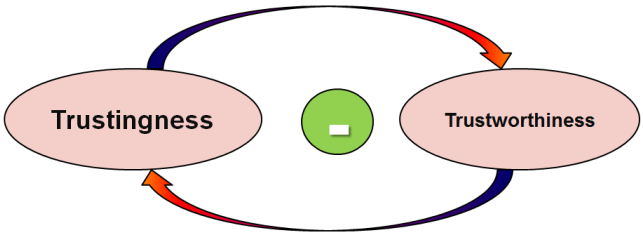


Figure 4.1: The two trust measures introduced in this study negatively reinforce each other.

[1] http://snap.stanford.edu/data/soc-Epinions1.html

This research has the following contributions:

- A pair of complementary global trust measures for a social trust network
- A classification system of networks based on risk involved to create links in a network
- Modeling involvement (by a Zipf distribution) and negative feedback property using a decay function
  - Error Bounds of the decay function

The chapter is organized as follows: Section 2 provides a survey of the research done in the broad area of computational trust and more specifically in the area of scoring trust in social networks. The next section sets up the problem of computing trust scores. Next the proposed approach is discussed followed by a section on algorithmic analysis. Finally a section on experiments and results is presented and eventually the chapter is wrapped by putting forth the conclusions and future work.

4.3 Related Work

Iterative matrix algorithms to compute abstract scores [66] of entities have been around for a long time. It was introduced in the field of marketing research to compute the influence of a product in its market segment [67], [68]. Kamakura *et. al.* [68] proposed a pair of measures driven by a product‘s market share. The proposed measures are **competitive clout & vulnerability**, referring to the impact the product has on the market shares of its competitors and a product‘s susceptibility to have their market share change as a result of price change of a competitor respectively. These measures complement each other. A successful product is expected to have a high competitive clout and low vulnerability. In [68], the authors have computed scores of four products for various segments of the society. The primary assumption in the chapter was that all four products impacted each other in the market. The analysis lacked an underlying network structure.

Graph theoretic models were used in the late 1990s by PageRank [36] and HITS [35] to rank nodes in a network. The context of usage of both algorithms was to aid the then

fledgling state of web search. The algorithms popularized web search. Consequently a host of other methods building upon either of the two methods [69] or proposing entirely new algorithms have been introduced [70]. Like [68], HITS introduced two complementary measures as the means for finding "authoritative" web pages on a given web search query. Given a web search query, HITS creates a network of web pages and calculates the *hubs* and *authority* scores for all pages in the network. However, both in HITS and in the preceding work by Kamakura *et. al.* [68], the pair of measures proposed reinforces each other positively, i.e., increase in one measure of a node leads to an increase in the other measure of its neighbors. The situation completely changes when the measures negatively reinforce each other. The algorithm behaves very differently and the convergence of the iterative matrix algorithm does not follow the patterns shown in [35].

During the last decade various researchers have tried to assign trust scores [71], [72], [73], [33] to nodes in a network to accomplish various tasks. Trust scores can be defined as scores that an algorithm puts on a node in a trust network based on various structural aspects of the node. Eigentrust [72] proposes to rate trust scores of peers in a P2P network. These scores help an ordinary user in the network to identify the trustworthy peers and initiate content download from them. This introduces policing inside a P2P network and discourages the dishonest peers to spread malicious and/or bogus content. Eigentrust, like Pagerank [36] calculates a single score for each node in the network. The study refers to it as the trust score. This score is calculated as a function of the trust/distrust votes a node gets based on the quality of content it is sharing with its peers. However, in this algorithm, one's reputation does not play a part in the weight of the node's trust vote.

Researchers have proposed measures to rank *bias* and *deserve* of a node in a network [73]. Like HITS, the research uses an iterative matrix algorithm to calculate bias and deserve of nodes which reinforce each other. In computation of deserve, according to [73], the authors rely only on the quality of inlinks. Here the paper tries to ascertain the reputation of a node inside a network. In failing to capture quantity along with quality of nodes, the paper fails to utilize the full potential of the whole network structure.

4.4 Computing Trust Scores in a Network

"Trust/Reputation Scores" in a social network is defined as a single or a set of scores that is assigned to each actor in the network representing his level of trust in the network. Researchers [72] have used single scores in network to depict the reputation of a node in the network. In this work, instead of assigning a single score, a pair of scores have been assigned to each actor in the network. These scores are referred to as "trustingness" and "trustworthiness" of actors in a network.

Primarily the objective of this research is twofold:

- Quantification of the abstract concept of trust in social networks. This is done in a 2 phase process.
  - Use a survey to determine trusting-decision involvement or simply **involvement** of social networks,
  - Use involvement and negative feedback property in trust to quantify it into 2 scores.
- Application of the trust scores to solve social network prediction class of problems.

The problem of calculating trust scores in a network is stated as follows:

4.4.1 Problem Definitions

The problem of finding trust scores in a social network can be defined as follows:

Given a directed social network $G = V, E$ and its trusting-decision involvement, where each edge is denoted by $e(uv) \in E$ represents a directed edge between source node $u \in V$ and destination node $v \in V$ and may or may not have weight $w(uv)$. In terms of a trust network, the edge $e(u, v)$ represents node u trusting node v.

In its current form, the problem outputs 2 scores (*trustingness* and *trustworthiness*) for each actor in the network. The primary constraints for the problem is that the sum of trustworthiness of all nodes in the network equals 1 and the sum of trustingness scores of all nodes in the network sums up to 1.

$$\sum_{v \in V} Trustingness(v) = 1 \tag{4.1}$$

$$\sum_{v \in V} Trustworthiness(v) = 1 \tag{4.2}$$

Problem Statement: Trust Scores

Formally, the problem of finding trustingness and trustworthiness in a network can be defined as follows:

Given:

1. A directed network $G < V, E >$ where V represents a set of all actors (nodes, used interchangeably) in the network and E represents the set of all edges in the network,
2. A convergence value δ,
3. Involvement of the directed social network.

Compute:

- For each actor $v \in V$
    1. Trustingness
    2. Trustworthiness

Constraint:

1. Sum of trustingness of all actors = 1 as represented in equation A.1.
2. Sum of trustworthiness of all actors = 1 as represented in equation A.2.

Readers can change the normalization criteria to suit his own requirements.

As can be seen in the problem statement, one of the input for computing trust scores in a network is the level of "**involvement** of a given network". Next, a problem statement for computing involvement in a directed social network will be formalized.

Problem Statement: Involvement of a Social Network

Laurent and Kapferer in [74] defined involvement of a network as the amount of loss a node stands to incur when it creates a wrong link. Jain and Srinivasan in [75] provided evidence of more and more social scientists accepting Laurent's definition of involvement in networks. In other words involvement is defined as the potential loss of an actor in a network for creating a wrong link.

Applying the definition of trusting-decision involvement henceforth referred to as involvement, from the consumer behavior research (e.g., Jain and Srinivasan [75] and Laurent and Kapferer in [74]), involvement of a social network is defined as an actor‘s perceived importance of trusting-decisions within the network and perceived risk or loss in case of wrong decisions.

In this research, a user survey is used to determine the involvement score of different networks, because involvement is a concept inherently perceived by network users. The survey was designed where respondents were provided with description of different social networks and asked a series of 7-point scale questions. These questions assessed the respondents‘ perceived importance of making decisions to link or not to link to others within the network along with perceived risks involved.

4.5 Approach: Trust Score Calculation

4.5.1 Calculation of Involvement of a Social Network

The concept of involvement in a social network was introduced in the previous section. Laurent and Kapferer in [74] defines involvement in social networks as the potential risk an actor takes when he is creating a link in the network. In a highly involved network, an actor stands to lose much more compared to a low involved network when he creates a potentially wrong link.

The survey questionnaire was designed by adopting well-established involvement measurement with a series of 7-point scales. In the survey questionnaire a detailed description of each social network was provided. Primarily, what each node in the particular social network refers to was listed and also what each link in the network represents. To measure involvement, the respondents were asked 5 questions which

are considered proxies for involvement. The questions measured the potential risk of creating a wrong link in the network and the perceived risks associated with the networks. The survey questionnaire for a sample network (Twitter Retweet network) had the following questions:

1. In deciding to retweet in Twitter, would you say that:
   - I would not care at all whose message I retweet $\Longrightarrow$ I would care a great deal
2. Do you think that the users of Twitter would be all very alike or all very different in terms of their trustworthiness for your retweeting?
3. In making your decision to retweet someone in Twitter, how concerned would you be about the outcome of your choice?
4. How do you feel about the potential risk of retweeting a wrong person in Twitter?
5. How important would it be for you to make a right choice of retweeting a person in Twitter?

Survey Sample

The survey, approved by IRB[2] , was administered within a sample of undergraduate students in a US mid-western research university. The sample of respondents was recruited from Computer Science and Mass Communication departments and a total of 123 participants took part in the survey. Out of the 123, 69 were male, 53 female and 1 person did not wish to disclose his/her sex. The median age of the respondents was 20 with the majority of the respondents being Caucasians (105 out of 123).

Survey Compilation

The most important part of any survey is to check the consistency of the answers provided by the respondents. It ensures that the responses provided are not random. To check the consistency of the responses, Cronbach's α test is used. Cronbach's α, which is a co-efficient of internal consistency is commonly used as an estimate of reliability of

[2] Institutional Review Board

survey measurement. In a sum of K components ($K-$items or $K-$testlets), Cronbach's α is defined as [76]

$$alpha = \frac{K}{K-1}\left(1 - \frac{\sum_{i=1} K\sigma^2_{Y_i}}{\sigma^2_X}\right) \tag{4.3}$$

where σ^2_X is the variance of the observed total test scores, and $\sigma^2_{Y_i}$ the variance of component i for the current sample of persons. In other words Cronbach's α measures the ratio of sum of variances of responses of every single test question to the variance of the entire test. A lower variance indicates that the respondents are consistent in their answers. It also indicates that there is very little randomness involved on the respondents' part while answering the questions.Cronbach's α score is presented for each surveyed social network in table 5.2 on page 100.

Once an acceptable consistency score (Cronbach's α) was reached for a network, the responses of the survey from a 7 point scale was normalized to a score between 0 and 1. For each question in the survey, all respondents' answers were compiled across all networks. To normalize the scores between 0 and 1, 0 was considered as mean - 1 standard deviation (say L) and 1 as mean + 1 standard deviation (say U). To calculate the normalized score of each network the mean score (out of 7), say M was calculated, and was normalized using the equation $\frac{M-L}{U-L}$. The involvement score for each surveyed social network is shown in table 5.2 on page 100.

4.5.2 Trust Scores: Basic Concepts

In this section the concepts of Trust Scores in a social network $G(V, E)$, where V denotes the set of all edges and E denotes the set of all edges, is revisited:

In Function

The function *in* of a node $in(v)$ where $v \in V$ is defined as a set of nodes which are the source nodes for all the incoming edges of node v.

Out Function

The function *out* of a node $out(v)$ where $v \in V$ is defined as a set of nodes which are the destination nodes for all the outgoing edges of node v.

Trustingness

Trustingness of an actor is defined as his propensity to trust others in the network. A higher trustingness score necessarily implies that the actor has a high propensity to trust others in the network.

Trustworthiness

Trustworthiness, true to its dictionary meaning, defines how trustworthy an actor is. Like trustingness score, a higher trustworthiness score means the actor is a highly trustworthy person in the network.

Trust Score: Properties

The primary property leveraged to calculate trust scores is the negative feedback property of trust. The concept of negative feedback in trust can be well understood using the example network provided in figure 4.2.

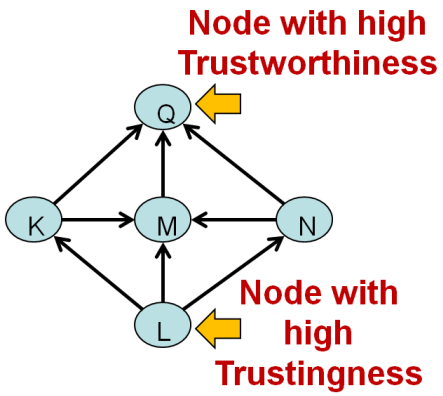


Figure 4.2: An example network where edges indicate source trusting destination.

In figure 4.2, there are nodes (say L), which has a high propensity to trust other nodes. L trusts almost all nodes in the network, except 1 (Node Q). Thus it can be seen that the L will accord trust to almost anyone in the network which should decrease the weight of its trust vote compared to a node like M which accords its trust very selectively.

Conversely, it can be seen that node Q is a highly trusted node. A high number of nodes in the network trust it. Moreover the nodes that trust (K, M, N) it in turn trusts a very selective amount of other nodes which makes their(K, M, N's) votes more valuable compared to L's.

Using the negative feedback property described above, it can be said that a higher **trustingness** score contributes to the trustworthiness of its neighbors to a lower degree. And a higher **trustworthiness** score is a result of lots of neighbors having low trustingness scores. In a variably weighted network, a person's trustingness depends on the edge weights of the outgoing edges. An actor's trustingness is given by:

$$trustingness(v) = \sum_{\forall x \in out(v)} \left(\frac{w(v,x)}{1 + trustworthiness(x)} \right) \tag{4.4}$$

Equation 4.4 suggests that the trustingness depends on three factors:

- Trustworthiness of the destination nodes
- Number of outgoing links
- Edge weight of each outgoing link

Similarly an actor's trustworthiness is given by:

$$trustworthiness(u) = \sum_{\forall x \in in(u)} \left(\frac{w(x,u)}{1 + trustingness(x)} \right) \tag{4.5}$$

Equation 4.5 suggests that the trustworthiness function depends on three factors:

- Trustingness of the source node
- Number of incoming links
- Edge weight of each incoming link

Trust Scores: Hypothesis

While introducing involvement in page no., 69, it was mentioned that involvement of a social network will be used to calculate trust scores in a network. To understand the usage, the use of decay function in trust score calculation will be introduced. As seen in

the last section, an increase in the value of 1 score (say trustworthiness) inversely impacts the 2nd score (trustingness) of its neighbors. Decay function helps in characterizing this property which can be seen in the formaliztions of the scores presented in equations 4.4 & 4.5.

Level of involvement of a network is defined as the amount of risk involved in making a wrong link in the network. Higher the risk in a social network(i.e., higher the involvement score), higher should be the effect of a neighbor's trustingness on the calculation of a node's trustworthiness and vice versa. Using this hypothesis and a Zipf distribution, it is claimed that trustworthiness is inversely proportional to sum of involvement exponent of neighboring nodes' trustingness and trustingness is inversely proportional to sum of involvement exponent of neighboring nodes' trustworthiness. Thus the equations in 4.4 & 4.5 get transformed into equations 4.6 & 4.7 respectively.

$$ti(v) = \sum_{\forall x \in out(v)} \left(\frac{w(v,x)}{(1 + (tw(x))^s} \right) \tag{4.6}$$

$$tw(u) = \sum_{\forall x \in in(u)} \left(\frac{w(x,u)}{1 + (ti(x))^s} \right) \tag{4.7}$$

where $ti(v)$ is trustingness of node v, $tw(v)$ is trustworthiness of node v and s is the involvement score of the given network.

To understand the use of involvement score, let us hypothetically consider 2 networks, one with involvement score(s) of 0 and other with $s = 1$. In the network with $s = 0$, there is no risk involved with creating wrong links. Thus calculation of trustworthiness should not be affected by neighbor's trustingness. Substituting $s = 0$ in equation 4.7, the equation transforms to $tw(u) = \sum_{\forall x \in in(u)} \left(\frac{w(x,u)}{1+(ti(x))^0} \right)$. The modified equation shows that making $s = 0$ converts the trustingness of neighbors into 1 and thus trustworthiness becomes a function of the quantity of connections and quality. Conversely in a network with $s = 1$, the risk involved in creating a wrong link becomes very high. Thus, while calculating trustworthiness, trustingness of neighbors should highly affect the trustworthiness score of a node. Equation 4.7 gets modified into $tw(u) = \sum_{\forall x \in in(u)} \left(\frac{w(x,u)}{1+(ti(x))^1} \right)$. In this case, since the proposed approach is using $(ti(x))^1$, trustingness of neighbors is entirely affecting the a node's trustworthiness scores.

4.6 TSM: Algorithm to Compute Trust Scores

In this section the TSM (**T**rust scores in **S**ocial **M**edia) algorithm is presented to measure trustingness and trustworthiness of actors in the network. The phrases **TSM** and **Trust Scores** are used interchangeably throughout the study and is intended to mean the same algorithm (Algorithm 1) whose description is provided below.

4.6.1 Algorithm

The last section discusses how equations 4.6 and 4.7 reinforce each other. They are mutually recursive in the sense that trustingness of an actor is dependent on the trustworthiness of its neighbors and vice versa. An iterative matrix convergence algorithm is used to solve the problem of finding trust scores in a network. Although TSM is a HITS-like algorithm, but is considerably different from the original idea of HITS proposed by [35]. HITS is a 2-score scoring system where only the scores from previous iteration affect calculation of scores in the current iteration. However is TSM, the porposed approach uses a 3^{rd} external factor called involvement (of social networks) as introduced in equations 4.6 and 4.7 via a Zipf function. Moreover HITS uses a positive feedback loop and thereby uses simple matrix manipulation for its convergence. On the other hand TSM leverages the negative feedback property of trust thereby changing the whole concept of HITS. The idea of convergence is not straight forward since to the best of the author's knowledge no such convergence proof exists for algorithms leveraging negative feedback.

TSM is an iterative matrix convergence algorithm/ It takes the equations presented in equations 4.6 & 4.7 and iterates over it as shown in algorithm 1. Trustingness takes the trustworthiness scores of all $out(v)$ from the previous iteration. The same is applicable for the trustworthiness calculation. The modified equations are given below:

$$ti_i(v) = \sum_{\forall x \in out(v)} \left(\frac{w(v,x)}{(1 + (tw_{i-1}(x))^s} \right) \tag{4.8}$$

$$tw_i(u) = \sum_{\forall x \in in(u)} \left(\frac{w(x,u)}{1 + (ti_{i-1}(x))^s} \right) \tag{4.9}$$

To the best of the author's knowledge, this is the first time, an iterative convergence algorithm has been used to model an abstract human emotion.

Data: 1) a directed graph $G = (V, E)$ consisting of vertices and edges with or without weights, and,
2) maximum number of permitted iterations k, and/or,
3) Difference of scores between 2 iteration, δ.

Result: A set of 2 trust scores(trustingness(ti), trustworthiness(tw)) $\forall v \in V$.

Initialize all $v \in V$ to $(1, 1)$;

for *(*$i = 1$*;* $max(max(|ti_i(v) - ti_{i-1}(v)|), max(|tw_i(v) - tw_{i-1}(v)|)) < \delta$ *or* $i \leq k$*; ++* i*)* **do**

  for *each node* $v \in V$ **do**

    update scores of each vertex using scores from last iteration;

    $ti'_i(v) = \sum_{\forall x \in out(v)} \left(\frac{w(v,x)}{(1+(tw_{i-1}(x))^s} \right)$;

    $out(v)$ = set of all vertices which are destination vertex of all outgoing edges from v;

  end

  for *each node* $v \in V$ **do**

    $tw'_i(u) = \sum_{\forall x \in in(u)} \left(\frac{w(x,u)}{1+(ti_{i-1}(x))^s} \right)$;

    $in(v)$ = set of all vertices which are source vertex of all incoming edges to v;

  end

  $ti_i = Normalize\,(ti'_i)$;

  $tw_i = Normalize\,(tw'_i)$;

end

Algorithm 1: Algorithm to calculate Trust Scores

TSM described in algorithm 1 takes a directed graph as input and asks the user for a convergence criteria or a maximum permitted number of iterations. In each iteration, for each node in the network, trustingness and trustworthiness is calculated using the equations presented in equations 4.8 and 4.9. Once the measures are calculated for each node in the network, the scores are normalized by adhering to the normalization constraints the user chooses to use. At the start of each iteration a convergence criterion is checked. If the difference of values between the last two iterations is less than the user defined δ parameter or if the total number of iterations is greater that user defined k, the algorithm converges.

4.6.2 Algorithmic Complexity

TSM is an iterative algorithm presented in algorithm 1. For each iteration, the trustingness and trustworthiness scores need to be calculated for each actor in the network. Trustingness is calculated using equation 4.6. Assuming $|E|$ to be the total number of edges present in the network, calculating trustingness requires time in the order of $O(|E|)$ since each edge has to be computed once in the use of equation 4.8. Similarly computation of trustworthiness according to equation 4.9 requires computation of each edge once. In case of the calculation of trustingness, the trustworthiness of the destination node of each edge is used whereas in case of calculation of trustworthiness, trustingness score of the source node of each node is used. The time required for the calculation of trustingness and trustworthiness in each iteration is of the order $O(|E|)$. The rest of the operations like normalization, etc., in each iteration is of the order of nodes present in the network $O(|V|)$. Since it is assumed that the number of nodes in the network is less than the number of edges ($|E| > |V|$), it can be concluded that the running time for TSM is $k \times O(|E|) = O(k \times |E|)$.

4.7 Algorithmic Analysis

4.7.1 Rate of Convergence

In this section, it will be shown that the maximum deviation of the value of any of the two trust measures for an actor is bound by an inverse exponential function dependent on

the number of iterations. For this proof, it is assumed that a slightly different version of the measures introduced in equations 4.6 and 4.7, which belongs to same family of decay function(inverse decay function). In the proof the Laplacian correction is disregarded, and it is assumed that $s = 1$. Moreover the proof assumes a local normalization factor.

The proof in this section will use trustworthiness to prove the error bounds. Since trustingness and trustworthiness are mirror images of one another, interested readers can use trustingness to prove the same.

It is assumed in this proof that an actor can reach his "true" trustworthiness scores in infinite iterations. Thus, the difference between the actual trustworthiness of a node $v \in V$ and the trustworthiness of v at an iteration i is given by $|trustworthinss_{\infty}(v) - trustworthiness_i(v)|$. In this section it will be proved this value is bound by an inverse exponential function on the number of iterations. For the sake of convenience, the term trustworthiness will be replaced by an abbreviation **tw** and trustineness by **ti**. Thus, an equation like 4.5 will look like.

$$tw(u) = \frac{1}{2|in(v)|} \sum_{\forall x \in in(u)} \left(\frac{w(x,u)}{ti(x)} \right) \tag{4.10}$$

Moreover, trustingness will be shortened to **ti**. Although this proof uses an infinite iteration to reach the "true" trustworthiness score of an actor, it will be shown using this proof that the family of function TSM belongs to converges in a small number of iterations.

Theorem 1 *The difference between the values of trustworthiness between any two iterations is less than equal to* 1*.*

Proof: According to equation A.2, it is shown that the value of trustworthiness score will always be normalized between 0 and 1. Thus the difference of trustworthiness between any two consecutive iterations can not be more than 1.

Lemma 2 ***Prove:***

$$\sum_t \frac{1}{xy} \leq \sum_t \frac{1}{x} \times \sum_t \frac{1}{y} \tag{4.11}$$

where $t \in \mathbb{N}$

Proof: Expanding R.H.S

$$\left(\frac{1}{x_1}+\frac{1}{x_2}+\ldots+\frac{1}{x_t}\right)\times\left(\frac{1}{y_1}+\frac{1}{y_2}+\ldots+\frac{1}{y_t}\right)$$

$$=\frac{1}{x_1y_1}+\frac{1}{x_2y_2}+\ldots+\frac{1}{x_ty_t}+\frac{1}{x_1y_2}+\ldots+\frac{1}{x_1y_t}+\ldots+\frac{1}{x_ty_{t-1}}$$

$$=\sum_t\frac{1}{xy}+\frac{1}{x_1y_2}+\ldots+\frac{1}{x_1y_t}+\ldots+\frac{1}{x_ty_{t-1}}$$

Since $\frac{1}{x_1y_2}+\ldots+\frac{1}{x_1y_t}+\ldots+\frac{1}{x_ty_{t-1}}$ is non-negative R.H.S $\geq$ L.H.S.

Theorem 3 *The difference between the trustworthiness score of an actor at an iteration i and the "true" trustworthiness score of the actor is bounded by an inverse exponential function having a function in the order of iteration i.*

$$|tw_\infty(v)-tw_i(v)|\leq\frac{1}{2^i} \tag{4.12}$$

Proof Mathematical induction is used for the proof From equation 4.9 and replacing the Laplacing correction,

$$tw_{i+1}(v)=\frac{1}{2|in(v)|}\sum_{\forall x\in in(v)}\left(\frac{w(x,v)}{ti_{i+1}(x)}\right) \tag{4.13}$$

Now substituting the value of $ti_{i+1}(x)$ from equation 4.8 in equation 4.13 :

$$tw_{i+1}(v)=\frac{1}{2|in(v)|}\sum_{\forall x\in in(v)}\left(\frac{w(x,v)}{\frac{1}{|out(x)|}\sum\limits_{\forall y\in out(x)}\left(\frac{w(xy)}{tw_i(y)}\right)}\right) \tag{4.14}$$

Now substituting the value of trustworthiness from equation 4.14 in the following equations

$$tw_\infty(v)=\frac{1}{2|in(v)|}\sum_{\forall x\in in(v)}\left(\frac{w(x,v)}{\frac{1}{|out(x)|}\sum\limits_{\forall y\in out(x)}\left(\frac{w(xy)}{tw_\infty(y)}\right)}\right) \tag{4.15}$$

$$tw_1(v)=\frac{1}{2|in(v)|}\sum_{\forall x\in in(v)}\left(\frac{w(x,v)}{\frac{1}{|out(x)|}\sum\limits_{\forall y\in out(x)}\left(\frac{w(xy)}{tw_0(y)}\right)}\right) \tag{4.16}$$

Basis Step The proof is for $i = 1$

$$|tw_\infty(v) - tw_1(v)| = \frac{1}{2|in(v)|} \sum_{\forall x \in in(v)} \left(\frac{w(x,v)}{ti_\infty(x)} - \frac{w(x,v)}{ti_0(x)} \right) \tag{4.17}$$

$$\begin{aligned}
&|tw_\infty(v) - tw_1(v)| \\
&= \frac{1}{2|in(v)|} \sum_{\forall x \in in(v)} \left(\frac{w(x,v)}{\frac{1}{|out(x)|} \sum_{\forall y \in out(x)} \left(\frac{w(xy)}{tw_\infty(y)} \right)} - \right. \\
&\left. \frac{w(x,v)}{\frac{1}{|out(x)|} \sum_{\forall y \in out(x)} \left(\frac{w(xy)}{tw_0(y)} \right)} \right)
\end{aligned}$$

[Assuming an equally weighted network $w(a,b) = 1$]

$$\begin{aligned}
&= \frac{1}{2|in(v)|} \sum_{\forall x \in in(v)} \left(\frac{1}{\frac{1}{|out(x)|} \sum_{\forall y \in out(x)} \left(\frac{1}{tw_\infty(y)} \right)} - \right. \\
&\left. \frac{1}{\frac{1}{|out(x)|} \sum_{\forall y \in out(x)} \left(\frac{1}{tw_0(y)} \right)} \right) \\
&= \frac{1}{2|in(v)|} \sum_{\forall x \in in(v)} \\
&\left(\frac{\frac{1}{|out(x)|} \left[\sum_{\forall y \in out(x)} \frac{1}{tw_0(y)} - \sum_{\forall y \in out(x)} \frac{1}{tw_\infty(y)} \right]}{\frac{1}{|out(x)|} \left[\sum_{\forall y \in out(x)} \left(\frac{1}{tw_0(y)} \right) \times \sum_{\forall y \in out(x)} \left(\frac{1}{tw_\infty(y)} \right) \right]} \right)
\end{aligned}$$

From Theorem 1,

(4.18)

Induction Step Assuming that the maximum deviation of trustworthiness at the i^{th} iteration is bounded by $\frac{1}{2^i}$, prove that the maximum deviation in the $(i+1)^{th}$ iteration is bounded by $\frac{1}{2^{i+1}}$. The deviation in the $(i+1)^{th}$ iteration is given by

$$\begin{aligned} & |tw_\infty(v) - tw_{i+1}(v)| \\ & = \frac{1}{2|in(v)|} \sum_{\forall x \in in(v)} \left(\frac{w(x,v)}{\frac{1}{|out(x)|} \sum_{\forall y \in out(x)} \left(\frac{w(xy)}{tw_\infty(y)} \right)} - \right. \\ & \left. \frac{w(x,v)}{\frac{1}{|out(x)|} \sum_{\forall y \in out(x)} \left(\frac{w(xy)}{tw_i(y)} \right)} \right) \end{aligned} \tag{4.19}$$

Replacing tw_1 with tw_{i+1} and tw_0 with tw_i in the set of equations represented in 4.18,

$$\begin{aligned} & |tw_\infty(v) - tw_{i+1}(v)| \\ & \leq \frac{1}{2|in(v)|} \sum_{\forall x \in in(v)} \left(\frac{1}{out(x)} \sum_{\forall y \in out(x)} |i| \right) \end{aligned} \tag{4.20}$$

$$i = |max0, tw_\infty(v) - max0, tw_i(v)|$$

In this case either $i = 0$ or $i = |tw_\infty(v) - tw_i(v)|$. Thus $|i| \leq |tw_\infty(v) - tw_i(v)| \leq \frac{1}{2^t}$ Therefore replacing values in equation 4.20

$$|tw_\infty(v) - tw_{i+1}(v)| \leq \frac{1}{2^{i+1}} \tag{4.21}$$

Hence proved

In this proof, it is shown that for a similar family of function to the one used for calculating trust scores, the difference between trustworthiness score of an actor at an iteration "*i*" and his *true* trustworthiness score is bounded by an inverse function having a function in the order of iteration "*i*".

4.7.2 Convergence

To prove the convergence of TSM, the error bounds need to be used. The final convergence proof will be in similar vein to other iterative algorithms like Bias and Deserve [73] and SIMRANK [77].

Let us assume that TSM converges at iteration k. Now using the rate of convergence function, proved in the previous section, it is needed to prove that the trustworthiness of the node at iteration k is less than the maximum deviation set at iteration k.,

Trustworthiness convergence can be defined as $|tw_{\infty}(v) - tw_k(v)| \leq \epsilon$ where $\epsilon \to 0$. Thus it can be concluded $k > log\left(\frac{1}{\epsilon}\right)$.

4.8 Experimental Evaluation & Results

4.8.1 Datasets

Several real life datasets publicly and privately available are used in this work. The publicly available datasets used in this chapter are Epinions dataset [33], Slashdot dataset [34], StackOverflow dataset and Twitter retweet dataset. The raw data for the StackOverflow dataset was downloaded from Stackexchange archive[3] . When an user marks another user's question as "favorite", it is considered that a trust link has formed between the 2. The Twitter retweet dataset was made available. Retweeting in Twitter, refers to the fact that the retweeter trusts the original tweeter‘s message. Thus, in this dataset, retweeting is used as a proxy for trust.

The EQ2 dataset [22] used in this chapter is a gaming log from EverQuest II developed by Sony Online Entertainment. The data is collected over a 35 week period and is completely anonymized. The data spans over various servers to make sure all types of activities are captured. A summary of the network used is presented for a better comprehension of the dataset.

Trust is an abstract concept. Proxies of trust are required to be identified which can be scientifically mapped to the original concept of trust. In this dataset, housing access in EverQuest II is used as a proxy for trust.

[3] `https://archive.org/details/stackexchange`

EQ II: House Network

Every character in the game is entitled to buy in-game houses [78]. Houses serve as a refuge to store in-game virtual items amassed in the game. Thus, from the perspective of in-game wealth, houses are vitally important to their owners. In EverQuest II, a player can "trust" his in-game *friend* and allow the person access to his/her house. The friend can view, interact and move objects in and out of these houses. When an owner of a certain house (henceforth referred to as the truster) grants access of his house to an in-game *friend* (henceforth referred to as the trustee), an edge in the housing network is introduced. Granting access to one's house to a different character in the game involves risk since the trustee can "steal" objects from the house that the owner (truster) has put effort to amass.

Table 4.1: Snapshot of the datasets used

Datasets	Epinions	Slashdot	EverQuestII
Nodes	75879	77360	78125
Edges	508837	905468	180256
Nodes, edges in WCC	1.0, 1.0	1.0, 1.0	0.8,0.9
Nodes, edges in SCC	0.425, 0.872	0.909, 0.981	0.41, 0.78

A snapshot of the datasets are provided in table 5.2 along with Cronbach's α score & involvement score for the datasets. The details of the public datasets are taken from the Stanford Network Analysis Project's dataset collection [4] .

Table 4.2: Snapshot of the datasets used

Datasets	Nodes	Edges	Cronbach's α	Involvement Score
Stack Overflow	134523	1597888	0.858	0.552
EverQuest II	63918	128048	0.841	0.811
Epinions	75879	508837	0.785	0.667
SlashDot	77360	905468	0.858	0.552
Twitter	1012012	9013252	0.767	0.359

[4] `http://snap.stanford.edu/data/`

4.9 Experiments

Various analysis comparing trust scores with in-degree and out-degree of all the social network datasets were performed. Moreover, comparison of the score distribution of the proposed score with HITS [35] is shown in Appendix A.

In this section the results of using trust scores to predict trust formation in social networks is presented. The proposed approach is compared to state of the art trust scoring algorithms. Trustingness and trustworthiness of those individuals are high in a network who have a high propensity to trust other and who are trustworthy by nature respectively. It is hypothesized that if a person with high trustingness is geodesically close to a person with high trustworthiness, a trust link should form. To exploit this hypothesis, a trust prediction task comparing the proposed approach is conceived along with state of the art trust scoring approaches like Bias-Deserve by [73] and HITS by [35]. A better trust link prediction shows that the proposed idea conforms to the original human idea of trust.

The prediction task was setup as a binary classification task where the attributes were the scores from scoring algorithms. This would make the comparison a fair comparison. The positive instances in the dataset are the ones where the actual links were present whereas negative instances were the ones where the links were absent but the nodes were within a geodesic distance of 3 [79]. The algorithms used were Bias-Deserve, HITS and 2 variations of the Trust Scores (TSM) algorithm. The "Adjusted Trust Scores" is the variant of TSM algorithm which uses the involvement parameter. "Trust Scores" is the variant without the involvement scores factored in. The results are tabulated in figure 4.3.

The second set of experiment demonstrated here is "Precision @ K" charts and "Precision-Recall" curves. For these experiments, all nodes were ranked with highest trustworthiness-trustingness product pairs within a certain geodesic distance. For "precision @ K", the ratio of number is checked for links actually formed to K. For precision-recall for each precision, the recall is calculated and plotted. By definition, a person with high trustingness should form a link with a highly trustworthy person. Thus, if in reality this is happening in a real social network, it provides a validity of the proposed concept.

4.9.1 Results

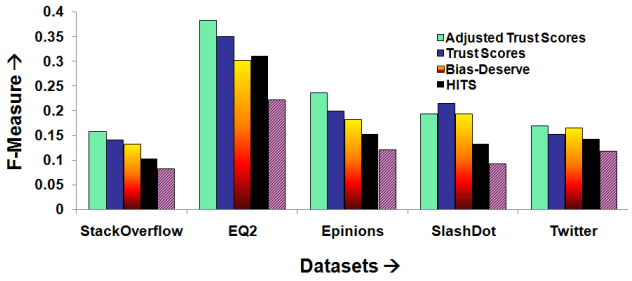
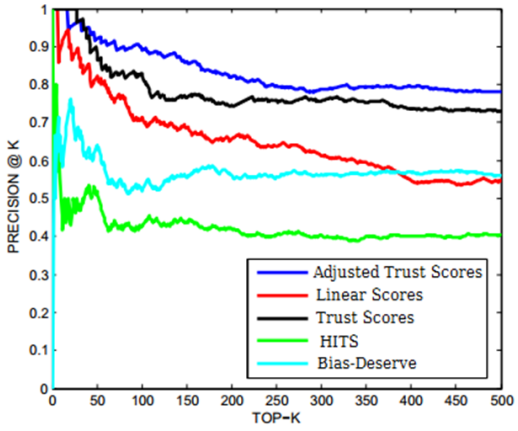
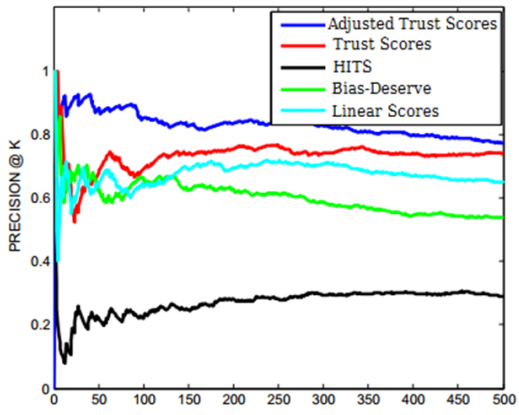


Figure 4.3: F-measures of trust prediction by various algorithms.

The results of trust prediction can be found in figure 4.3. It can be seen that for majority of networks Adjusted Trust Scores performs better than all the other algorithms. In SlashDot it is found that Trust Scores performing better than Adjusted Trust Scores. SlashDot is a computer application related news bulletin and the respondents from Mass Communication might not be the typical users of this website. Is is suspected that the involvement score determined for this specific network might not be a true reflection of the risk involved in creating a wrong link in this network.

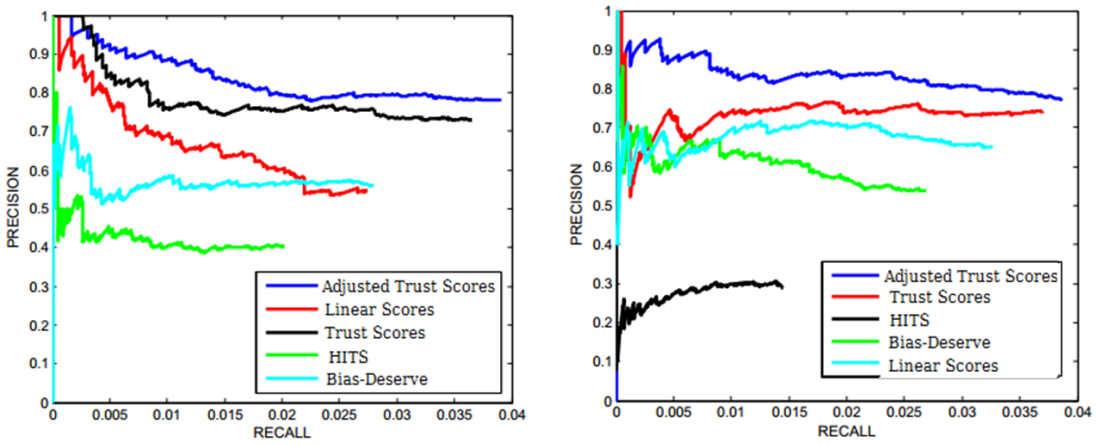


(a) Precision at K chart for EverQuestII dataset.



(b) Precision at K chart for Epinions dataset.

Figure 4.4: Precision at K chart for various datasets



(a) Precision Recall curves for EverQuestII dataset.

(b) Precision Recall curves for Epinions dataset.

Figure 4.5: Precision Recall curves for various datasets

In both figures 4.4(b) & 4.4(a), it is found that Adjusted Trust Scores have higher precision than other techniques for all values of K. Adding the results demonstrated in figures 4.5(b) & 4.5(a), it can be claimed that trust definitely is governed by the ideas of trustingness and trustworthiness and a use of a negative feedback loop is the best way to capture the essence in a social network.

4.10 Case Study

4.10.1 Identification of Rumor Spreading Paths in Hurricane Sandy Tweets

For the purpose of performing an acid test of the proposed approach, an experimental case study was done using tweets collected on Hurricane Sandy. The dataset consists of sequences of re-tweets by various Twitter users who tweeted on the topic of Hurricane Sandy. The primary aim of the case study was to find the trustworthy sources and also to identify sources who have a propensity to spread misinformation and rumor.

Using re-tweet sequences from the dataset, a network was formed where re-tweets was considered a proxy of trust. When a person re-tweets, he trusts the judgment of the original poster and thereby spreads his view. The proposed approach assigned very low

trustworthiness score to the posters who were later identified as the source of spreading misinformation. The image shown in figure 4.6 was a well known rumor which got circulated during the aftermath of Hurricane Sandy [80]. The user who tweeted the image in figure 4.6 had a very low trustworthiness score.

Moreover the algorithm was also able to identify the highly trustworthy sources. Typically the sources that had the highest trustworthy scores were the reputed media houses like CNN [5] and FoxNews [6] .

Figure 4.6: Example of rumor spread during the aftermath of Sandy hurricane

4.11 Conclusion & Future Work

Assigning scores to actors in a trust network is crucial for several applications. This chapter introduces two complementary concepts of trust, trustingness and trustworthiness which have negative feedback properties. Unlike EigenTrust, while calculating the trust scores of actors, this approach takes into account not only the incoming links, but also the reputation of the truster. The algorithm proposed in this chapter is efficient since it has an algorithm complexity of $O(k \times |E|)$ and is shown to converge very quickly.

[5] http://www.cnn.com/
[6] http://www.foxnews.com/

The set of experiments performed show that the measures are analogous to other ranking theorems. Finally, a case study on real life data shows the effectiveness of these measures. On one hand they were capable of identifying actors spreading rumors and on the other hand they were also capable of identifying reputed organizations which are trusted highly by the community.

Currently the proposed measures are capable of measuring trust scores of actors in a network. There are various applications which can benefit from a score like this. Primarily an application is proposed where these scores can be used for stopping rumor spread in networks or conversely can be used for maximizing influence in networks thereby helping in applications like viral marketing. Since it is seen that the probability of formation links for a highly trusting person is high, the rumor spreaders generally use these channels to spread rumors in the networks. These vulnerability increases when a highly trusting person becomes very trustworthy in the network. The in-links to the node become highly vulnerable and become potential for rumor spread in the network. An application of trust scores can be used to identify these nodes which can stop the flow of rumor spread.

Moreover, instead of stopping the flow, identifying these nodes may also create potential influence maximizers and influence flow paths in the network which can be leveraged by agencies to virally market their products.

Chapter 5

Identification of Vulnerable Paths in Social Networks

5.1 Overview

The future work of The last chapter alluded to the fact that the concept of negatively reinforced trust scores can be used in various application areas. In some cases the application of trust scores is proposed to start a new way the problem was viewed and in other cases it builds over the existing solutions and provides newer insights into both the problem space and also improves the accuracy of the current state-of-the-art algorithms. This chapter discusses a couple of these application areas in great details and will gloss over a few potential areas which can will benefit with the introduction of trust scores.

5.2 Introduction

The primary objective of every entrepreneur is to reach the widest possible audience for her products. Across the ages the techniques used has varied which has always been defined by the technologies available at the time. The development of televisions, satellites and radios have changed the way a product was advertised. After the emergence of internet, online social network and online social media, the space of advertisements have seen a radical shift. The advertisements not only include selling material/electronic

products but also services and most importantly ideas.

The new age of advertisements have dawned upon us, and more and more advertisers are trying to leverage the social media to sell their products in an already crowded market place. The primary motivation for the advertiser in a social domain is to catch the eye of his target consumer. A consumer will be enticed in ones product, only if they are excited with the product and/or if the recommendation comes from someone that the consumer "trusts" [81, 82]. For example in the Twitter network, this can equate to so someone more likely to try a product or an idea if it endorsed by another person whom this person follows or "retweets" regularly. Jansen *et. al.* in [81] shows that even in 2009 the companies were exploiting these ideas in the Twitter space to promote their products. The "word of mouth" advertisements is not a new cultural phenomena that has emerged in the age of online social networks. It has always been there [83]. A person is always more likely to use a product/service if a near and dear one (a person whom he "trusts") recommends.

With the dawn of the online social networks and social media these dynamics have changed. These online social networks have erased geographic boundaries and have enabled people to voice their opinions to hundreds of thousands of users (if not millions). Thus a product can become viral easily since the information exchange can occur easily compared to what it was even two decades ago. Advertisers have taken note of this phenomena and are trying to leverage this while selling their products.

This chapter proposes a technique that identifies those actors in a network who are in optimal topographical position (in a social network) to aid in viral marketing.

With the aid of trust scores calculated for each node in the network, introduced in the last chapter, this chapter furthers the investigation. Here instead of calculating scores for each node in the network, scores for each edge is calculated. The scores are a function of trustingness of one node and trustworthiness of the other node. The algorithm takes as an input a threshold also known as "vulnerability threshold(α)". The edges that have a score greater than α are considered "vulnerable" edges. This algorithm looks for "long" vulnerable paths in the network. The hypothesis behind the "vulnerable" paths is that information flow happens very easily along these paths in a social network.

To measure the accuracy of this study, a prediction task is proposed at the end of

the study which compares the trust path predictions and compares it against scores and paths generated by state of the art trust scoring algorithms.

5.2.1 Motivation

Social network analysis have always helped viral marketing in its new avatar. It has helped understand who can be potential targets and who are the ones that should be targeting based on the influence and trust they have/share. Kleinbergś early study on Influence propagation and later studies have proved it so. But the biggest problems in these studies is that how to calculate the influence and how to exploit it.

Moreover weighing edges in a social networks have other applications too. For example in Kempeś seminal work on influence propagation [84], the edges weighed which were referred to as influence probability were modeled randomly. Over the years researchers have used various algorithms to model edge weights. But none of them to the best of the author's knowledge have leveraged the social network to put influence propagation weights in the edges. Subbian in[85] have used extraneous information like number of retweets and number of papers co-authored as weights to these edges. This research uses only the network topographical structure to find these edges weights. This study can also use those extraneous information that the other studies have used to calculate the information transition probabilities for the edges in a social network.

5.2.2 Contributions

This study makes the following contributions:

1. This study looks not only at dyadic relations in a network but paths. This study is able to leverage actors' influence over other actors whom they may not know in real life.

2. This study shows that there are a few actors in a network who can impact other person's decision to accept or reject a product by a series of actors present in the path. This is done using the concept of "vulnerable paths" in a network. This is the first study that exploits the notion of "vulnerable" paths in a social network to understand how influence propagates in a social network.

Figure 5.1: Companies who actively use online social viral marketing or are used extensively as a medium for the same.

- This study uses Trust Scores from previous chapter to identify the vulnerable paths through which the study proposes that influence will propagate.

3. The notion of calculating edge weights can also be used in other applications directly for influence propagation transition probabilities.

5.3 Related Works

The problem of finding influencers in the network is often studied as an influence maximization problem [84, 86, 87, 85]. The problem of influence maximization is finding the top-k nodes such that the average infection spread is maximized, under a specific influence propagation model. There are two popular choices for the influence propagation model, Independent Cascade (IC) and Linear Threshold (LT) [84]. All these related work assume edge propagation probabilities for the influence propagation model are given. The most popular choices for edge propagation probabilities are weighted

cascade model [84] or trivalency model [88].

These techniques assume that the infection probabilities are provided as an input to the social directed network. In other words these methods assume that somehow the probabilities with which a person A will trust/influence person B is provided to the algorithm. This becomes very tricky to estimate.

In this work instead of looking at finding the seeders in the network trust scores will find pathways which will help in flow of influence in the social network.

5.4 Trust Scores: A Brief Description

5.4.1 Computing Trust Scores in a Network

'Trust/Reputation Scores" in a social network is defined as a single or a set of scores that is assigned to each actor in the network representing his level of trust in the network. Researchers [72] have used single scores in network to depict the reputation of a node in the network. In this work, instead of assigning a single score, a pair of scores have been assigned to each actor in the network. These scores are known as "trustingness" and "trustworthiness" of actors in a network.

Calculating trust scores is a 2-step process:

- Use a survey to determine trusting-decision involvement or simply **involvement** of social networks,
- Use involvement and negative feedback property in trust to quantify it into 2 scores by exploiting the social network structure.

5.4.2 Calculating Involvement of a Social Network

Involvement of a social network is a user survey to determine the involvement score of a given social network. Since involvement of a social network is a concept inherently perceived by network users, a survey is designed where respondents were provided with description of different social networks and asked a series of 7-point scale questions. These questions assessed the respondents' perceived importance of making decisions to link or not to link to others within the network along with perceived risks involved thereby providing a normalized (between 0 and 1) score for the given network.

5.4.3 Basic Concepts

In Function

The function *in* of a node $in(v)$ where $v \in V$ is defined as a set of nodes which are the source nodes for all the incoming edges of node v.

Out Function

The function *out* of a node $out(v)$ where $v \in V$ is defined as a set of nodes which are the destination nodes for all the outgoing edges of node v.

Trustingness

Trustingness of an actor is defined as his propensity to trust others in the network. A higher trustingness score necessarily implies that the actor has a high propensity to trust others in the network.

Trustworthiness

Trustworthiness, true to its dictionary meaning, defines how trustworthy an actor is. Like trustingness score, a higher trustworthiness score means the actor is a highly trustworthy person in the network.

Trust Score: Properties

The primary property leveraged in this research to calculate trust scores is the negative feedback property of trust. Using the negative feedback property it can be postulated that a higher **trustingness** score contributes to the trustworthiness of its neighbors to a lower degree. And a higher **trustworthiness** score is a result of lots of neighbors having low trustingness scores. In a variably weighted network, a person‘s trustingness depends on the edge weights of the outgoing edges.

$$trustingness(v) = \sum_{\forall x \in out(v)} \left(\frac{w(v,x)}{1 + trustworthiness(x)} \right) \tag{5.1}$$

Similarly an actor's trustworthiness is given by:

$$trustworthiness(u) = \sum_{\forall x \in in(u)} \left(\frac{w(x,u)}{1 + trustingness(x)} \right) \tag{5.2}$$

5.4.4 Edge Score

Trustingness and trustworthiness of all the nodes can be calculated for a social network. Edge score of a directed edge $A- > B$ indicating A "trusts" B is defined as the product of the trustingness score of A with the trustworthiness score of B.

5.4.5 Vulnerable Edges

An edge which has an edge score $es \leq$ some vulnerability threshold is defined as a vulnerable edge.

5.4.6 Vulnerable Paths

In a social network a path is defined as a sequence of directed edges. A vulnerable path is defined as the sequence of edges in a social where each of the edge in the sequence is a vulnerable edge.

5.5 Assumptions

The primary assumption in this research is if an actor **A** "trusts" another actor **B** in a network, then B has some "influence" over A [37]. For example in a social network like Twitter if A "follows" B, A is more likely to retweet B's tweets. Thus in this research the reader needs to remember that the influence flow in the network is in opposite direction to the trust flow in the network.

5.6 Problem Statement

The problem of finding vulnerable paths in a social network can be defined as follows.

Given:

1. A directed network $G < V, E >$ where V represents a set of all actors (nodes, used interchangeably) in the network and E represents the set of all edges in the network,
2. A vulnerability threshold α,
3. Trust scores ti, tw $\forall v \in V$
4. A minimum path length β

Compute:

- find all vulnerable paths that are greater than equal to 2 (or β, when provided)

Objective:
Constraint:

- The edges should be subsequent to each other and there should be one direction of influence flow.
- The weight of each edge in the paths should be higher than α

5.7 Approach

This section provides a detailed analysis of the algorithm to find vulnerable paths in a social network. To understand this algorithm the reader needs to have a good understanding of the Trust Scores algorithm presented in Chapter 4 of this thesis. The primary assumption of the algorithm is that the algorithm is already provided with the trust scores for each node in the network.

5.7.1 Finding Vulnerable Paths

The first step in the algorithm of finding vulnerable paths is identifying vulnerable edges in the network. Given the vulnerability threshold α, find all the edges that have a vulnerability higher than α. Once the vulnerable edges are identified in the network, re draw the social network with only the "vulnerable" edges.

The network so drawn will be used to find paths greater than the provided path threshold β $G' = (V, E')$.

5.7.2 Algorithm to find "Vulnerable Paths"

This section will discuss the algorithm to find vulnerable paths of length *gen* from a modified network G'.

Data: 1) a directed graph $G' = (V, E')$ consisting of vertices and edges where edges represent vulnerability weights,
2)path length n.

Result: A master list of vulnerable paths $list_{paths}$

$list_{paths} =$ **for** $i = diameter(G'); i >= n; i--$ **do**
| $list_{paths} -> list_{paths} +$ algorithm 3(G', n)
end

Algorithm 2: Algorithm to find all vulnerable paths in a network

Algorithm 2 is a iterative algorithm which starts at diameter of the network and call algorithm 3 every time for decreasing values of i until the optional parameter of path length (n) is met. The output is stored in a master list which is the final result of the algorithm.

Algorithm to find paths of length n

Data: 1) a directed graph $G' = (V, E')$ consisting of vertices and edges where edges represent vulnerability weights,
2)path length x.

Result: A master list of vulnerable paths of length x, m_{list}

for *each available vertex* $v \in V$ **do**
  Set its value to $seen = 1$ ($seen(v) = 1$);
  if *number of vertices* $|path|$ *in the path equals the desired length* ($v' == x$) **then**
    Store path in master list ($m_{list} = path$)
  else
    Set "available vertices" to all unseen adjacent vertices($neighbor(v)$);
    Repeat from top
  end
  Remove the latest vertex and add it to the path ($path = path + v$);
  Un-select the vertex ($seen(v) = 0$)
end

Algorithm 3: Algorithm to find paths of length x in a network

The primary task of algorithm 2 is to compute all paths of a given length in a network. A detailed algorithm for that task is provided in algorithm 3.

Algorithm 3 takes in a modified network (only consisting of vulnerable edges) G' and an optional parameter x. This algorithm is a modular function to the algorithm 2. It starts iterating over all available vertex. For each iterating vertex a flag is set and the vertex is push to a current path. If the length of the path is of desired length x, the algorithm adds the path in to the set of paths of length x. If the condition is not satisfied, the algorithm sets all "available vertices" to all unseen vertices and repeats the algorithm. Next the vertex in question is added to the path and the vertex is unselected.

5.8 Experiments & Results

5.8.1 Datasets

Several real life datasets used in the last research are used in this research. The publicly available datasets used in this paper are Epinions dataset [33], Slashdot dataset [34], StackOverflow dataset and Twitter retweet dataset. The raw data for the StackOverflow dataset was downloaded from Stackexchange archive[1] . When an user marks another user's question as "favorite", it is considered that a trust link has formed between the 2. The Twitter retweet dataset was made available. Retweeting in Twitter, refers to the fact that the retweeter trusts the original tweeter‘s message. Thus, in this dataset, retweeting is used as a proxy for trust.

The EQ2 dataset [22] used in this paper is a gaming log from EverQuest II developed by Sony Online Entertainment. The data is collected over a 35 week period and is completely anonymized. The data spans over various servers to make sure all types of activities are captured. A summary of the network used is presented for a better comprehension of the dataset.

Trust is an abstract concept. Proxies of trust are required to be identified which can be scientifically mapped to the original concept of trust. In this dataset, to be housing access in EverQuest II is identified as a proxy for trust.

EQ II: House Network

Every character in the game is entitled to buy in-game houses [78]. Houses serve as a refuge to store in-game virtual items amassed in the game. Thus, from the perspective of in-game wealth, houses are vitally important to their owners. In EverQuest II, a player can "trust" his in-game *friend* and allow the person access to his/her house. The friend can view, interact and move objects in and out of these houses. When an owner of a certain house (henceforth referred to as the truster) grants access of his house to an in-game *friend* (henceforth referred to as the trustee), an edge in the housing network is introduced. Granting access to one‘s house to a different character in the game involves risk since the trustee can "steal" objects from the house that the owner (truster) has put effort to amass.

[1] `https://archive.org/details/stackexchange`

A snapshot of the datasets are provided in table 5.2 along with Cronbach's α score & involvement score for the datasets. The details of the public datasets are taken from the Stanford Network Analysis Project's dataset collection [2] .

Table 5.1: Snapshot of the datasets used

Datasets	Epinions	Slashdot	EverQuestII
Nodes	75879	77360	78125
Edges	508837	905468	180256
Nodes, edges in WCC	1.0, 1.0	1.0, 1.0	0.8,0.9
Nodes, edges in SCC	0.425, 0.872	0.909, 0.981	0.41, 0.78

Table 5.2: Snapshot of the datasets used

Datasets	Nodes	Edges	Involvement Score
Stack Overflow	134523	1597888	0.552
EverQuest II	63918	128048	0.811
Epinions	75879	508837	0.667
SlashDot	77360	905468	0.552
Twitter	1012012	9013252	0.359

5.8.2 Experiments

The primary motivation behind the experiments section is to prove the usefulness of the vulnerable paths in a social network. The first experiment provides a statistics on the number of vulnerable paths present in a network. One of the input to the "Finding Vulnerable Paths" algorithm is the vulnerability threshold α of a network. This part of the research experiments with various sets of vulnerability threshold in a network. The primary objective of this analysis is to figure out the various thresholds for various social networks.

It was already proved that trustingness and trustworthiness of those individuals are

[2] `http://snap.stanford.edu/data/`

high in a network who have a high propensity to trust other and who are trustworthy by nature respectively. The last study proved the hypothesis that if a person with high trustingness is geodesically close to a person with high trustworthiness, a trust link should form. The hypothesis used in this study is similar to the last one. Instead of predicting links in the network, this work predicts paths in the social network. Moreover this study looks at a neighborhood of $4-6$ hops. To exploit this hypothesis, the trust prediction task proposed, compares this approach with state of the art trust scoring approaches like Bias-Deserve by [73] and HITS by [35] used in the last study. A better trust path prediction should validate the fact that the proposed idea conforms finding vulnerable paths in a social network.

The prediction task was setup as a binary classification task where the attributes were the scores from scoring algorithms. This would make the comparison a fair comparison. The positive instances in the dataset are the ones where the actual links were present whereas negative instances were the ones where the links were absent but the nodes were within a geodesic distance of $4-6$. The algorithms used were Bias-Deserve, HITS and the Trust Scores algorithm. The results are tabulated in figure 4.3.

The second set of experiment demonstrated here is Precision @ K charts and Precision-Recall curves. For these experiments, all nodes with highest trustworthiness-trustingness are ranked product pairs within a certain geodesic distance. For precision @ K, the ratio of number of links actually formed to K is checked. For precision-recall for each precision, recall is plotted. By definition, a person with high trustingness should form a link with a highly trustworthy person. Thus, if in reality this is happening in a real social network, it provides the validity of the proposed concept.

5.8.3 Results

The results of analysis of investigating "Vulnerability Threshold" can be found in figures 5.2, 5.3 and 5.4 for the Epinions dataset. It can be seen in the figure 5.2 that as vulnerability threshold α increases, the amount of paths decreases, which is what one expects. The interesting takeaway from the figures is the fact that there are several knees of the curves. Based on the requirement of the reader, he can choose set an α for his study. Note setting a higher α puts more restrictions. But the paths so found are highly vulnerable. On the other hand setting a lower α results in the algorithm

including less vulnerable paths in the result set with a large number of candidate sets. Thus if the application demands a high recall choosing a higher α is desirable whereas an application demanding a higher precision should choose a lower vulnerability threshold.

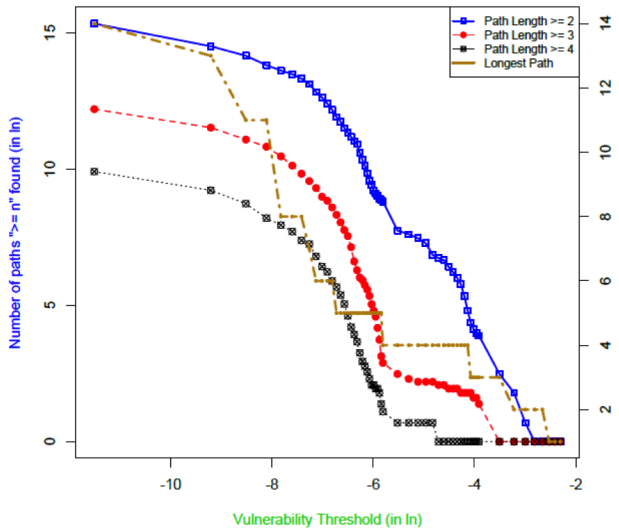


Figure 5.2: Comparison of number of paths $\geq n$ against vulnerability threshold in Epinions dataset. The right vertical axis in the chart represents the longest path for a specific vulnerability threshold.

Next figures 5.3 and 5.4 shows the accuracies of the various algorithms. **TS** in the legend refers to the Trust Score algorithm proposed in this study. **BD** refers to Bias-Deserve from [73] and **HITS** refers to seminal work by Jon Kleinberg [35]. Moreover the part "*Path Length ¿= 2*" refers to the fact that all paths are chosen whose length ≥ 2. The *x-axis* in the figure shows vulnerability threshold and the *y-axis* refers to the F-1 score for predictive accuracy. For a fair comparison, *TS Path Length ¿= 2* should be compared to *HITS Path Length ¿= 2* and *BD Path Length ¿= 2* and so on and so forth. It is evident that that the proposed vulnerable paths based on Trust Scores from last chapter consistently outperforms the other 2 state of the art trust scoring algorithms for various path lengths.

Figure 5.4 shows the Recall at K curve for top 200 nodes pairs in the algorithm. Here too Trust Scores consistently outperforms its peers. Recall at K is chosen since in this

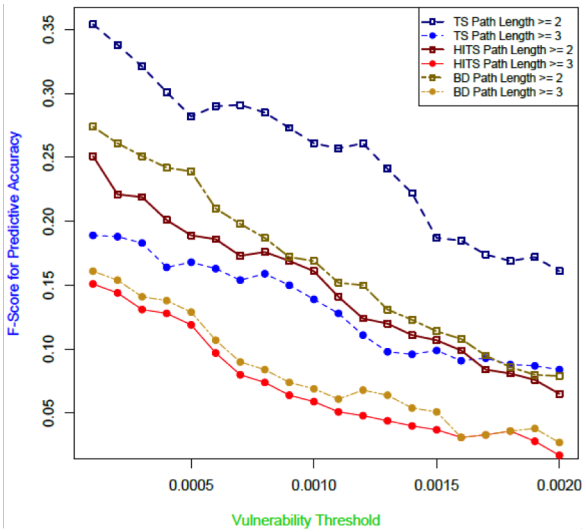


Figure 5.3: Accuracies of various algorithms in detecting trust paths in Epinions dataset.

application of viral marketing the user is most interested in finding highly vulnerable paths. He is more motivated by the fact that every vulnerable path should be discovered even that means he has to sift through multiple false positives. None of these paths should be left out. Thus recall is the perfect measure that the user intends to maximize in this particular application which lead to this experiment.

5.9 Conclusion & Discussion

This study proposes a new method to perform viral marketing using the concept of trust scores proposed in the last chapter. The primary idea in the study is to exploit the set of hypotheses that if a person with high trustingness is geodesically close to a person with high trustworthiness, a trust link should form. Moreover if a chain or path of such individuals can be found in a network, a vulnerable path can emerge through which influence can flow. Experiments on real life datasets show that the algorithm proposed in this study has greater accuracy in identifying these in a social network compared to other state of the art trust scoring algorithms like HITS and Bias-Deserve.

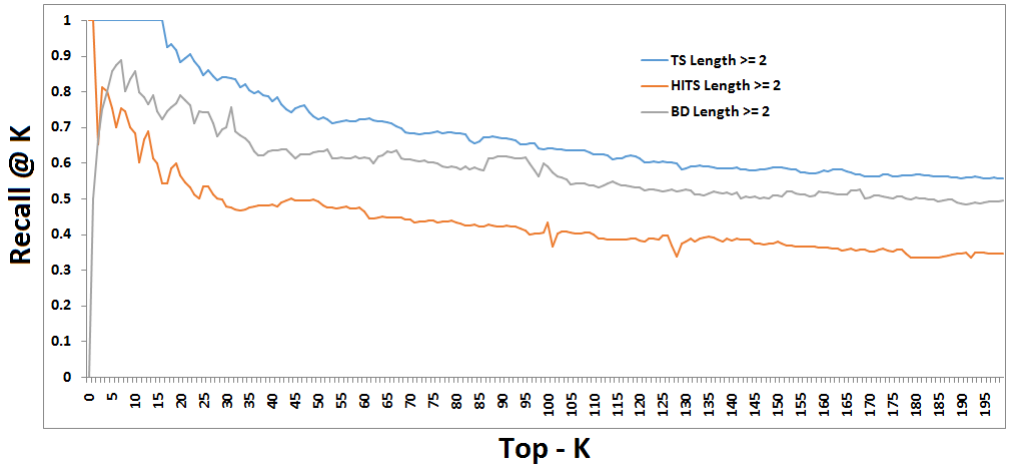


Figure 5.4: Recall at K curve for various trust scoring algorithms in the Epinions dataset

Chapter 6

Conclusion and Future Work

6.1 Conclusion

The literature on computation trust is huge. Trust has been studied in various disciplines and computer scientists have also joined in the effort. But the primary issue with these studies has been the problem of studying trust in isolation. As was seen throughout this thesis, trust has inter dependence on social interactions and studying or modeling trust without the important factor of social interactions will not yield in successful models. Moreover this study divides trust into various granularities and have identified the problems that plague each of these granularity. For the dyadic granularity this thesis has provided a entire state diagram of dyadic, identified various sub problems in it and have studied each of them, formation, reciprocation and revocation. For the global trust, this thesis has identified the social psychology that trust in humans tend to follow negative feedback property and has leveraged it to propose scores and use these scores in various applications like viral marketing.

6.2 Future Work

A number of problems were studied in this research. Answers to each of the sub-problem in this thesis have opened new avenues to study newer problems which were incomprehensible before this thesis. During the problem of identification of revocation, the model using metadata based dataset failed to elicit comparable predictive results.

A new avenue of study can ascertain the hypothesis stated in this thesis in section 3.9.3 and scientifically prove or disprove the hypothesis. Moreover figure 3.1 alludes to a problem of cascading trust revocation which could not be studied due to lack of data points. Cascading trust revocation can also be stated as "reciprocation of trust revocation".

Trust scores in social media **TSM** have various applications. One of the potential application discussed in this thesis is viral marketing. Viral marketing leverages the potential vulnerable paths in a social network and uses it to target consumers to sell one's products. Another application for the identifying vulnerable paths is to save a network from rumor spread. Misinformation just like information can spread in a network using the vulnerable paths. The idea of blocking rumor spread is to identify these vulnerable paths and inoculate the actors in these paths with actors having very low trustingness appearing in local neighborhood. Data to manipulate such hypothesis is hard to come by and thus this question was also not investigated in this thesis.

References

[1] Chandrima Sarkar, Sarah Cooley, and Jaideep Srivastava. Improved feature selection for hematopoietic cell transplantation outcome prediction using rank aggregation. In *FedCSIS*, pages 221–226, 2012.

[2] D Harrison McKnight and Norman L Chervany. The meanings of trust. 1996.

[3] D Harrison McKnight and Norman L Chervany. Trust and distrust definitions: One bite at a time. In *Trust in Cyber-societies*, pages 27–54. Springer, 2001.

[4] Niklas Luhmann. Trust and power. 1982.

[5] J David Lewis and Andrew Weigert. Trust as a social reality. *Social forces*, 63(4):967–985, 1985.

[6] Soyoen Cho, Jisu Huh, and Ronald J Faber. The influence of sender trust and advertiser trust on multistage effects of viral advertising. *Journal of advertising*, 43(1):100–114, 2014.

[7] Theodore M Porter. *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton University Press, 1996.

[8] Devon Johnson and Kent Grayson. Cognitive and affective trust in service relationships. *Journal of Business research*, 58(4):500–507, 2005.

[9] Jennifer Ann Golbeck. Computing and applying trust in web-based social networks. 2005.

[10] Stephen Paul Marsh. Formalising trust as a computational concept. 1994.

[11] Muhammad Aurangzeb Ahmad. *Computational trust in multiplayer online games.* University of Minnesota, 2012.

[12] C-N Ziegler and Georg Lausen. Spreading activation models for trust propagation. In *e-Technology, e-Commerce and e-Service, 2004. EEE'04. 2004 IEEE International Conference on*, pages 83–97. IEEE, 2004.

[13] Raph Levien. Attack resistant trust metrics. Technical report, 2004.

[14] Paolo Massa and Paolo Avesani. Trust-aware bootstrapping of recommender systems. In *ECAI Workshop on Recommender Systems*, pages 29–33. Citeseer, 2006.

[15] John O'Donovan and Barry Smyth. Trust in recommender systems. In *Proceedings of the 10th international conference on Intelligent user interfaces*, IUI '05, pages 167–174, New York, NY, USA, 2005. ACM.

[16] Jordi Sabater and Carles Sierra. Regret: A reputation model for gregarious societies. pages 61–69, 2001.

[17] Mark Witkowski, Er Artikis, and Jeremy Pitt. Experiments in building experiential trust in a society of objective-trust based agents. In *Trust in Cyber-societies, volume LNAI 2246*, pages 111–132. Springer-Verlag, 2001.

[18] Geraint Parry. Trust, distrust and consensus. *British journal of political science*, 6(02):129–142, 1976.

[19] Julian Rotter. A new scale for the measurement of interpersonal trust. *Journal of personality*, 1967.

[20] Dmitri Williams. The mapping principle, and a research framework for virtual worlds. *Communication Theory*, 20(4):451–470, 2010.

[21] Alice Leung, Will Dron, John P Hancock, Maitane Aguirre, Jon Purnell, Jiawei Han, Chi Wang, Jaideep Srivastava, Amogh Mahapatra, Atanu Roy, et al. Social patterns: Community detection using behavior-generated network datasets. In *Network Science Workshop (NSW), 2013 IEEE 2nd*, pages 82–89. IEEE, 2013.

[22] Atanu Roy, Zoheb Hassan Borbora, and Jaideep Srivastava. Socialization and trust formation: a mutual reinforcement? an exploratory analysis in an online virtual setting. In *ASONAM*, pages 653–660, 2013.

[23] Emma Burkitt Wright, Christopher Holcombe, and Peter Salmon. Doctors' communication of trust, care, and respect in breast cancer: qualitative study. *Bmj*, 328(7444):864, 2004.

[24] Muhammad Aurangzeb Ahmad, David A. Huffaker, Jing Wang, Jeffrey William Treem, Marshall Scott Poole, and Jaideep Srivastava. GTPA: A generative model for online mentor-apprentice networks. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2010, Atlanta, Georgia, USA, July 11-15, 2010*, 2010.

[25] Patricia M Doney and Joseph P Cannon. An examination of the nature of trust in buyer-seller relationships. *the Journal of Marketing*, pages 35–51, 1997.

[26] Roger C Mayer, James H Davis, and F David Schoorman. An integrative model of organizational trust. *Academy of management review*, 20(3):709–734, 1995.

[27] Muhammad Aurangzeb Ahmad. *Computational trust in Multiplayer Online Games*. PhD thesis, Department of Computer Science, University of Minnesota, Twin Cities, 2012.

[28] Muhammad Aurangzeb Ahmad, Iftekhar Ahmed, Jaideep Srivastava, and Marshall Scott Poole. Trust me, i'm an expert: Trust, homophily and expertise in mmos. In *PASSAT/SocialCom 2011, Privacy, Security, Risk and Trust (PASSAT), 2011 IEEE Third International Conference on and 2011 IEEE Third International Conference on Social Computing (SocialCom), Boston, MA, USA, 9-11 Oct., 2011*, pages 882–887, 2011.

[29] Zoheb Borbora, Muhammad A. Ahmad, Karen Zita Haigh, Jaideep Srivastava, and Zhen Wen. Exploration of robust features of trust across multiple social networks. In *SASO Workshops*, pages 27–32, 2011.

[30] Zoheb Hassan Borbora, Muhammad Aurangzeb Ahmad, Jehwan Oh, Karen Zita Haigh, Jaideep Srivastava, and Zhen Wen. Robust features of trust in social networks. *Social Netw. Analys. Mining*, 3(4):981–999, 2013.

[31] Ayush Singhal, Atanu Roy, and Jaideep Srivastava. Understanding co-evolution in large multi-relational social networks. In *Information Reuse and Integration (IRI), 2014 IEEE 15th International Conference on*, pages 733–740. IEEE, 2014.

[32] Atanu Roy, Muhammad Aurangzeb Ahmad, Chandrima Sarkar, Brian Keegan, and Jaideep Srivastava. The ones that got away: False negative estimation based approaches for gold farmer detection. In *SocialCom/PASSAT*, pages 328–337, 2012.

[33] Matthew Richardson, Rakesh Agrawal, and Pedro Domingos. Trust management for the semantic web. In *The Semantic Web-ISWC 2003*, pages 351–368. Springer, 2003.

[34] Jure Leskovec, Kevin J Lang, Anirban Dasgupta, and Michael W Mahoney. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics*, 6(1):29–123, 2009.

[35] Jon M Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, 46(5):604–632, 1999.

[36] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: bringing order to the web. 1999.

[37] Behnam Hajian and Tony White. On the interaction of influence and trust in social networks. In *Proceedings of the 1st Workshop on Incentives and Trust in E-Commerce*, pages 63–75, 2012.

[38] Ayush Singhal, Karthik Subbian, Jaideep Srivastava, Tamara G. Kolda, and Ali Pinar. Dynamics of trust reciprocation in multi-relational networks. In *Advances in Social Networks Analysis and Mining 2013, ASONAM '13, Niagara, ON, Canada - August 25 - 29, 2013*, pages 661–665, 2013.

[39] Jennifer Golbeck. *Computing with Social Trust*. Springer Publishing Company, Incorporated, 1st edition, 2008.

[40] Muhammad Aurangzeb Ahmad, Marshall Scott Poole, and Jaideep Srivastava. Network exchange in trust networks. In *Proceedings of the 2010 IEEE Second International Conference on Social Computing, SocialCom / IEEE International Conference on Privacy, Security, Risk and Trust, PASSAT 2010, Minneapolis, Minnesota, USA, August 20-22, 2010*, pages 341–346, 2010.

[41] Chandrima Sarkar and Jaideep Srivastava. Impact of density of lab data in ehr for prediction of potentially preventable events. In *Healthcare Informatics (ICHI), 2013 IEEE International Conference on*, pages 529–534. IEEE, 2013.

[42] Zoheb Borbora and Jaideep Srivastava. User behavior modelling approach for churn prediction in online games. In *SocialCom/PASSAT*, pages 51–60, 2012.

[43] Jehwan Oh, Zoheb Hassan Borbora, and Jaideep Srivastava. Automatic detection of compromised accounts in mmorpgs. In *2012 ASE International Conference on Social Informatics*, pages 222–227, 2012.

[44] Mohammad Al Hasan, Vineet Chaoji, Saeed Salem, and Mohammed J. Zaki. Link prediction using supervised learning. In *SDM06: Workshop on Link Analysis, Counter-terrorism and Security*, 2006.

[45] Chandrima Sarkar and Atanu Roy. Using gaussian measures for efficient constraint based clustering. *arXiv preprint arXiv:1411.3302*, 2014.

[46] David Liben-Nowell and Jon M. Kleinberg. The link prediction problem for social networks. In *CIKM*, pages 556–559, 2003.

[47] Leo Katz. A new status index derived from sociometric analysis. *Psychometrika*, 18(1):39–43, 1953.

[48] Jean C de Borda. Mémoire sur les élections au scrutin. 1781.

[49] Chandrima Sarkar, Sarah Cooley, and Jaideep Srivastava. Robust feature selection technique using rank aggregation. *Applied Artificial Intelligence*, 28(3):243–257, 2014.

[50] Chandrima Sarkar. *Improving Predictive Modeling in High Dimensional, Heterogeneous and Sparse Health Care Data*. PhD thesis, UNIVERSITY OF MINNESOTA, 2015.

[51] George Hripcsak and Adam S Rothschild. Agreement, the f-measure, and reliability in information retrieval. *Journal of the American Medical Informatics Association*, 12(3):296–298, 2005.

[52] Robin Dunbar. Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, 22(6):469–493, 1992.

[53] David Lazer, Alex Pentland, Lada Adamic, Sinan Aral, Albert-Lszl Barabsi, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, Tony Jebara, Gary King, Michael Macy, Deb Roy, and Marshall Van Alstyne. Computational social science. 323(5915):721–723, 2009.

[54] Nick Yee. Motivations for play in online games. *Cyberpsychology and Behavior*, 9(6):772–775, 2006.

[55] Edward Castronova. *Synthetic Worlds : The Business and Culture of Online Games*. University Of Chicago Press, 2005.

[56] Michael Szell and Stefan Thurner. Measuring social dynamics in a massive multiplayer online game. *Social Networks*, 32(4):313–329, 2010.

[57] Alvin W Gouldner. The norm of reciprocity: A preliminary statement. *American sociological review*, pages 161–178, 1960.

[58] Andreas Diekmann. The power of reciprocity fairness, reciprocity, and stakes in variants of the dictator game. *Journal of conflict resolution*, 48(4):487–505, 2004.

[59] Catherine A Bliss, Isabel M Kloumann, Kameron Decker Harris, Christopher M Danforth, and Peter Sheridan Dodds. Twitter reciprocal reply networks exhibit assortativity with respect to happiness. *Journal of Computational Science*, 3(5):388–397, 2012.

[60] Robert Eisenberger, Stephen Armeli, Barbara Rexwinkel, Patrick D Lynch, and Linda Rhoades. Reciprocation of perceived organizational support. *Journal of applied psychology*, 86(1):42, 2001.

[61] Marshall D Sahlins and Michael BANTON. On the sociology of primitive exchange in the relevance of models for social anthropology. 1965.

[62] Karen V. Hansen. The asking rules of reciprocity in networks of care for children. *Qualitative Sociology*, 27(4):421–437, 2004.

[63] Stephen Leider, Markus M. Mbius, Tanya Rosenblat, and Quoc-Anh Do. Directed altruism and enforced reciprocity in social networks. *The Quarterly Journal of Economics*, 124(4):1815–1851, 2009.

[64] Catherine A. Bliss, Isabel M. Kloumann, Kameron Decker Harris, Christopher M. Danforth, and Peter Sheridan Dodds. Twitter reciprocal reply networks exhibit assortativity with respect to happiness. *Journal of Computational Science*, 3(5):388–397, 2012.

[65] Muhammad Aurangzeb Ahmad, Brian Keegan, Atanu Roy, Dmitri Williams, Jaideep Srivastava, and Noshir S. Contractor. Guilt by association?: network based propagation approaches for gold farmer detection. In *ASONAM*, pages 121–126, 2013.

[66] Atanu Roy, Chandrima Sarkar, Rafal Angryk, et al. Using taxonomies to perform aggregated querying over imprecise data. In *Data Mining Workshops (ICDMW), 2010 IEEE International Conference on*, pages 989–996. IEEE, 2010.

[67] Lee G Cooper. Competitive maps: The structure underlying asymmetric cross elasticities. *Management Science*, 34(6):707–723, 1988.

[68] Wagner A Kamakura and Gary J Russell. A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research*, pages 379–390, 1989.

[69] Prasanna Desikan, Nishith Pathak, Jaideep Srivastava, and Vipin Kumar. Incremental page rank computation on evolving graphs. In *Special interest tracks and*

posters of the 14th international conference on World Wide Web, pages 1094–1095. ACM, 2005.

[70] Taher H Haveliwala. Topic-sensitive pagerank. In *Proceedings of the 11th international conference on World Wide Web*, pages 517–526. ACM, 2002.

[71] Donovan Artz and Yolanda Gil. A survey of trust in computer science and the semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(2):58–71, 2007.

[72] Sepandar D Kamvar, Mario T Schlosser, and Hector Garcia-Molina. The eigentrust algorithm for reputation management in p2p networks. In *Proceedings of the 12th international conference on World Wide Web*, pages 640–651. ACM, 2003.

[73] Abhinav Mishra and Arnab Bhattacharya. Finding the bias and prestige of nodes in networks based on trust scores. In *Proceedings of the 20th international conference on World wide web*, pages 567–576. ACM, 2011.

[74] Gilles Laurent and Jean-Noel Kapferer. Measuring consumer involvement profiles. *Journal of marketing research*, pages 41–53, 1985.

[75] Kapil Jain and Narasimhan Srinivasan. An empirical assessment of multiple operationalizations of involvement. *Advances in consumer research*, 17(1), 1990.

[76] Robert F DeVellis. *Scale development: Theory and applications*, volume 26. Sage Publications, 2011.

[77] Dmitry Lizorkin, Pavel Velikhov, Maxim Grinev, and Denis Turdakov. Accuracy estimate and optimization techniques for simrank computation. *Proceedings of the VLDB Endowment*, 1(1):422–433, 2008.

[78] EverQuest II: Housing Wiki. `http://eq2.wikia.com/wiki/Housing`.

[79] Salvatore Scellato, Anastasios Noulas, and Cecilia Mascolo. Exploiting place features in link prediction on location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1046–1054. ACM, 2011.

[80] Hurricane Sandy Photographs. `http://www.snopes.com/photos/natural/sandy.asp`, November 2012.

[81] Bernard J Jansen, Mimi Zhang, Kate Sobel, and Abdur Chowdury. Twitter power: Tweets as electronic word of mouth. *Journal of the American society for information science and technology*, 60(11):2169–2188, 2009.

[82] W Glynn Mangold and David J Faulds. Social media: The new hybrid element of the promotion mix. *Business horizons*, 52(4):357–365, 2009.

[83] Johan Arndt. *Word of mouth advertising: A review of the literature*. Advertising Research Foundation, 1967.

[84] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146. ACM, 2003.

[85] Karthik Subbian. *Scalable Analysis of Information Flows in Networks*. PhD thesis, University of Minnesota, 2014.

[86] Masahiro Kimura and Kazumi Saito. Tractable models for information diffusion in social networks. In *Knowledge Discovery in Databases: PKDD 2006*, pages 259–271. Springer, 2006.

[87] Jure Leskovec, Mary McGlohon, Christos Faloutsos, Natalie S Glance, and Matthew Hurst. Patterns of cascading behavior in large blog graphs. In *SDM*, volume 7, pages 551–556. SIAM, 2007.

[88] Wei Chen, Chi Wang, and Yajun Wang. Scalable influence maximization for prevalent viral marketing in large-scale social networks. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1029–1038. ACM, 2010.

Appendix A

Trustingness & Trustworthiness: A Pair of Complementary Trust Measures in a Social Network

A.1 Experimental Evaluation

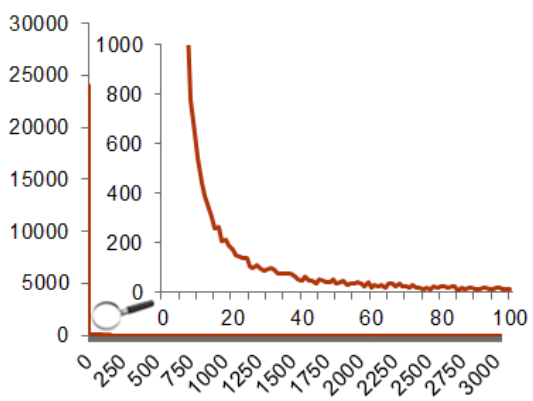
Throughtout this section, trustworthiness has been compared against indegree and authority score [35] whereas trustingness has been compared to outdegree and hub score [35] since the concepts can be considered analogous. Instead of using equations A.1 and A.2 for normalization, a different scheme is used where by the scores are normalized between 0 and 1.

$$\sum_{v \in V} Trustingness(v) = 1 \tag{A.1}$$

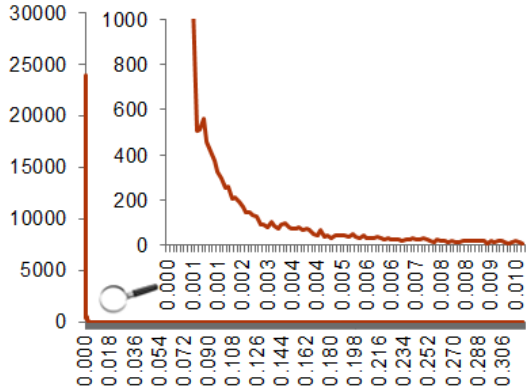
$$\sum_{v \in V} Trustworthiness(v) = 1 \tag{A.2}$$

A.1.1 Analysis of indegree and trustworthiness distribution

The first set of experiments are performed to analyze the distribution of indegree and the trustworthiness scores of all actors across the network. Figure A.1 shows the different graphs for the distribution of indegree and trustworthiness in various networks.



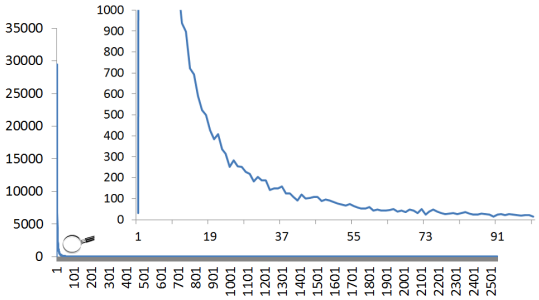
(a) Distribution of in degree for the Epinions dataset



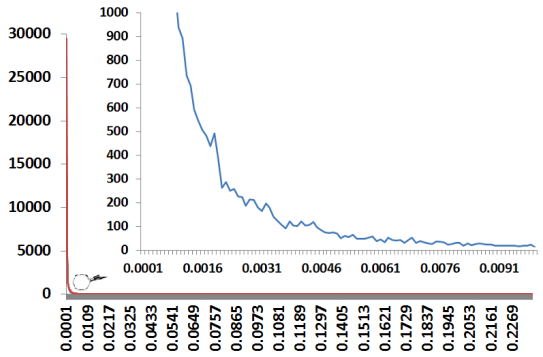
(b) Distribution of trustworthiness for the Epinions dataset

Figure A.1: Distribution of Trustworthiness and Indegree versus Frequency in Epinions dataset.

Figures A.1 & A.2 is a plot of trustworthiness/in degree versus frequency. The plot suggests the frequency of actors having a specific trustworthiness/in degree. The plot in the inset the figures in A.1(a), A.1(b), A.2(a) & A.2(b) show a magnified version of the original plot.



(a) Distribution of in degree for the SlashDot dataset



(b) Distribution of trustworthiness for the SlashDot dataset

Figure A.2: Distribution of Trustworthiness and Indegree versus Frequency in SlashDot dataset.

An analysis of figures A.1 & A.2, show that the distribution for trustworthiness is not smooth. The Epinions dataset has a number of disconnected components which lead to the irregular distribution of trustworthiness in figure A.1(b). Moreover, it can be

seen that there is a concentration of values at the lower range. Analysis of the SlashDot dataset demonstrated in figure A.2 also shows similar results. The figures show that the concept of trustworthiness can be compared to the concept of indegree, but is evident that both are not the same.

A.1.2 Analysis of trustingness versus trustworthiness distribution

This section compares the trustingness score of each actor versus their trustworthiness score. The x-axis in the sub figures represent an actor's trustingness score whereas the y-axis represents his trustworthiness score. Figure A.3 shows the distribution of trustingness versus trustworthiness for each actor in the various networks.

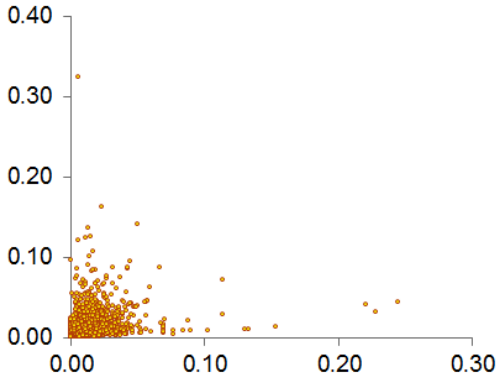
Analysis of all three datasets show the fact that majority of actors in the network tend to have low trustworthiness and trustingness scores. This is understandable since most of the actors in these networks are not hyper-active. In Epinions dataset, the actors tend to have higher trustingnessing score compared to trustworthiness score. As mentioned earlier in the paper, it is easy to have a high trustingness score whereas achieving a high trustworthiness score is tougher. In Epinions, actors trust each other's judgment for rating products and movies. An actor stands to lose less in terms of money and time if he trusts a wrong person. On the other hand, in case of EverQuest II housing networks, an actor stands to lose everything if he misplaces his trust. Thus, there are very few actors having high trustworthiness score in the network. Even trustingness scores are low too as is evident in figure A.3(b). SlashDot is a peculiar case in which the trustingness score and the trustworthy score are positively co-related with each other as can be seen in figure A.3(c).

A.1.3 Comparison with HITS

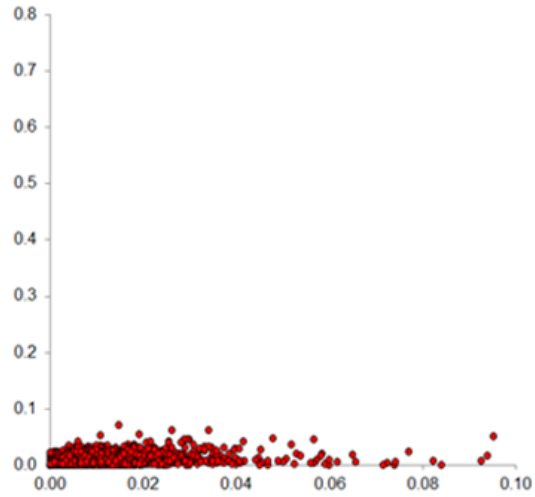
This experiment looks at the distribution of scores produced by a well known iterative scoring algorithm HITS [35]. Authority scores are considered analogous to trustworthiness scores whereas hub scores are considered analogous to trustingness scores. A distribution of hubs versus authorities is performed for the two publicly available dataset.

Comparison of figure A.4(a) with A.3(a) and figure A.4(b) with A.3(c) demonstrates

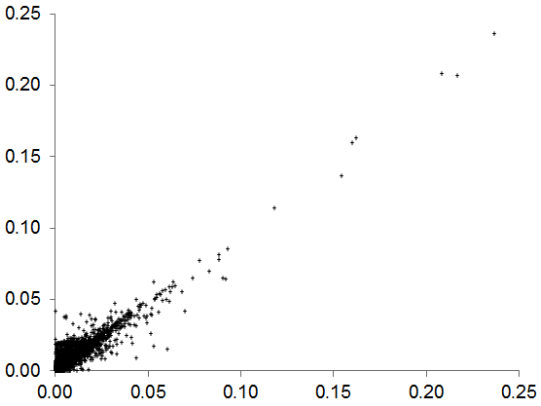
a similar trend in both the scoring algorithms. SlashDot dataset has a positive co-relation between the two measures whereas the Epinions dataset lacks so. On a close inspection it can be seen that the proposed approach leads to a higher variability in trust scores. This is because of the model that has been proposed. The proposed model takes into account both the quality and quantity of inlinks and outlinks and this results in trust scores across a greater range.



(a) Distribution of trustingness versus trustworthiness for the Epinions dataset

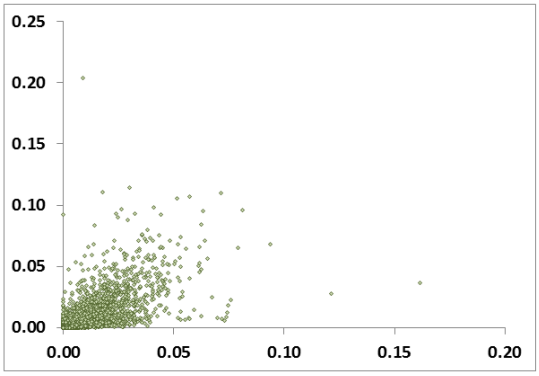


(b) Distribution of trustingness versus trustworthiness for the EverQuest II dataset

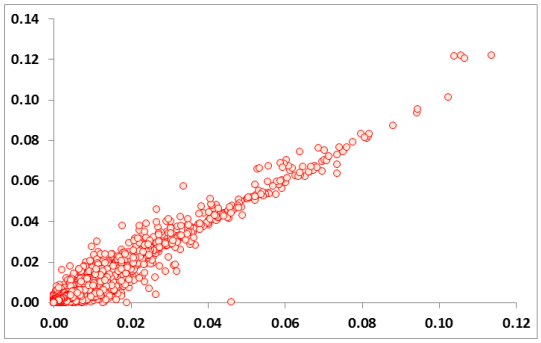


(c) Distribution of trustingness versus trustworthiness for the SlashDot dataset

Figure A.3: Distribution of trustingness versus trustworthiness for each actor in various networks



(a) Distribution of hubs versus authority scores for the Epinions dataset



(b) Distribution of hubs versus authority scores for the SlashDot dataset

Figure A.4: Distribution of hubs versus authority scores for each actor in various networks