Graduate School ETD Form 9
(Revised 12/07)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Pawat Chomphoosang

Entitled
Trust Management of Social Network in Heath Care

For the degree of Master of Science

Is approved by the final examining committee:

Arjan Durresi
Chair

Rajeev R. Raje

Yao Liang

To the best of my knowledge and as understood by the student in the *Research Integrity and Copyright Disclaimer (Graduate School Form 20)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): Arjan Durresi

Approved by: Shiaofen Fang 11/06/2012
Head of the Graduate Program Date

TRUST MANAGEMENT OF SOCIAL NETWORK IN HEALTH CARE

A Thesis

Submitted to the Faculty

of

Purdue University

by

Pawat Chomphoosang

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

May 2013

Purdue University

Indianapolis, Indiana

ACKNOWLEDGEMENTS

The work described in this thesis has been accomplished due to the assistance and support of many people to whom I would like to express my utmost gratitude. I would like to thank my research advisor, Dr. Arjan Durresi, for his encouragement and support as well as his invaluable advice during the thesis. Also, thanks to Dr. Rajeev R. Raje, Dr. Yao Liang, and Dr. Mohammad Al Hasan who have reviewed this thesis and have given me many good advises to improve the quality. Without the assistance of them, I could not accomplish the work. I am indebted to staff members of Department of Computer and Information Science for providing suggestions, assistance, and especially friendship which greatly supported me in my work. I would like to express my appreciation to my friends, especially Ping Zhang, Danar Widyantoro and Yefeng Ruan who have helped, either directly or indirectly, to stimulate my thought processes in this work. I would like to thank my family for their continual encouragement and patient during the time of study.

TABLE OF CONTENTS

Page

Page

LIST OF FIGURES

Figure Page

Figure Page

ABSTRACT

Chomphoosang, Pawat. M.S., Purdue University, May 2013. Trust Management of Social Network in Health Care. Major Professor: Arjarn Durresi.

The reliability of information in health social network sites (*HSNS*) is an imperative concern since false information can cause tremendous damage to health consumers. In this thesis, we introduce a trust framework which captures both human trust level and its uncertainty, and also present advantages of using the trust framework to intensify the dependability of *HSNS*, namely filtering information, increasing the efficiency of pharmacy marketing, and modeling how to monitor reliability of health information. Several experiments which were conducted on real health social networks validate the applicability of the trust framework in the real scenarios.

CHAPTER 1. INTRODUCTION

1.1 Introduction

There are more than twenty thousand health-related sites available on the Internet and over 62% of Americans as estimated by [1] have been influenced by the health information provided on news websites and the Internet, whereas 13% received the information from their physicians. Additionally, one study [2] shows that 87% of Internet users who look for health information believe that the information they read online about health is reliable, while another study [3] revealed that less than half of the medical information available online has been reviewed by medical experts and only 20% of Internet users verify the information by visiting authoritative websites such as CDC and FDA. As Health Social Networking Sites (*HSNS*) have emerged as a platform for disseminating and sharing of health-related information, people tend to rely on it before making healthcare decisions, such as choosing health care providers, determining a course of treatment and managing their health risks The work of [4] points out that the complex nature of *HSNS* has some unique challenges for both health consumers and service providers.

First, the health information is considered as highly sensitive information. Without deliberate consideration, the consumers may receive misleading information which may cause them severe damage. There are examples of misleading information written by [5].

Second, as health service providers, their reputation can be attacked by malicious users or honest users due to unethical competition or poor service. The report [6] describes that many physicians got negative reviews and ratings from review websites, and it’s unclear for viewers whether or not reviews and ratings are real. One possible solution is for the providers to attempt to eliminate the negative reviews. They may pay the owners of those sites to eliminate bad reviews or instead find someone to write good reviews to hide the negative reviews. As a result, both health consumers and service providers should be aware of several possible threats, including spreading disinformation, distributed denial of service, distorted advertisement and many others in the future. As in all systems dealing with information, *HSNS* will be successfully used if and only if it could provide reliability of information with a certain level of information security. Hence, the concept of trust will come into the picture.

1.2 Trust Framework

The trust framework [7] was developed based on the similarities between human trust operations and physical measurements. It consists of trust metrics and management methods to aggregate trust, which are based on measurement theory and guided by psychology and intuitive thinking. In general, the framework introduces two metrics, named *m* and *c*, both of which represent an interrelationship between nodes. *m* presents how one node, say Alice, evaluates the trustworthiness of another node, say Bob. Meanwhile, *c* represents how Alice is certain about the *m* opinion. We elaborate the theories and the framework further in Chapter 4. In this thesis, our purpose is to apply the trust framework to enable both individuals and system administrators to fulfill utilization of *HSNS* through the following functionalities.

First, individuals and administrators can use the framework for information filtering. If individuals use *m* and *c* metrics, the metrics can be a tool to assist the users whether information sources are reliable or not. Suppose, the consumer is looking for opinions about drug *A*, s/he is querying on his or her *HSNS*. Suppose there are many other users sharing both positive and negative opinions. S/he can use the trust transitive and aggregation equations to compute *m* and *c*, which are the indicators to discern the reliable information from the unreliable. The sources with low *c* are eliminated; meanwhile the sources with high *c* are being considered. In any case, if *m* opinions among sources of high *c* are similar, the consumer will gain more confidence(*c*) in the opinion. However, if *m* opinions among the sources are dissimilar, the consumer will lower *c*. This probably leads the consumer to acquire more information or the closed knowledge opinion leader (*KOL*), such as physicians or health experts, to regain *c*.

Second, administrators can also use the framework to improve optimized marketing tools. The existing tools aim to find a group of users who influence the greatest population in the network. One approach is to find a group of users who receive the most number of reviews and consider them as high influencers. Nonetheless, a number of reviews (only direct trust pointing to a user) is easy to generate. This technique is vulnerable to attackers. With the framework, we use both trust transitive and aggregation models in computing trust relations among users so-called *Trust Power*. It is a good indicator for improving the health marketing tools. A user with a higher score of *Trust Power* implies the higher power of influence to other nodes. We also note that a user who has a lot of direct trust relation does not necessarily have high *Trust Power*. After considering *Trust Power,* it is hard for malicious nodes to attack the system. Administrators can also use the framework to analyze the reliability of each information source. Sources that have high *Trust Power* are considered as reliable sources, while sources with low *Trust Power* are eliminated.

Third, administrators can also exploit the framework assist in monitoring reliability of a public opinion. Suppose *KOL* expresses an opinion about an object. The opinion probably makes an influence on his or her followers. As we mentioned *KOL* earlier, if many *KOLs* express opinions which are similar about the object, many followers who trust those *KOLs* will agree upon the consensus, and therefore the combined *Trust Power* of the object will be high. In other words, the reliable level of the particular object becomes high. Meanwhile, in case many *KOLs* express dissimilar opinions about the object, the confidence for their followers will be increasing, and consequently the combined *Trust Power* will be compromised. This indicates the low level of reliability for a particular object. Because of this, it is best for administrators to integrate the framework for monitoring the reliability of health products.

Fourth, we also compare the performance of our framework with another work [28] in two aspects: Robustness to attackers and identification of influencers. Based on the result, our framework outperforms the previous work.

1.3 Organization of this thesis

This thesis is organized as follows; we review possible sources where patients seek for information in Chapter 2. In Chapter 3, we explain possible issues in *HSNS*. In Chapter 4, we introduce a theoretical background of trust framework. Furthermore, we present the experiments and analysis that demonstrate that our methodology is applicable in the real world in Chapter 5. We compare the performance of our framework with the other framework in Chapter 6. In Chapter 7, we review related work in this domain. In Chapter 8, we present the conclusion and future work.

CHAPTER 2. SOURCES OF INFORMATION

Health consumers today tend to find health information on the Internet and then visit physicians. Therefore, there are several sources of health information online that health consumers reply on. We categorized them into the following four major services:

2.1 Health Web Portals

Health web portals are sources that provide health information which have been developed to educate patients. Patients can seek health information on them. For example, *www.webmd.com* is a very reliable source. Readers are more likely to trust its content as being developed by medical experts (*KOLs*). In the websites, patients cannot interact as much as web 2.0. As a result, trust evaluation is based on the portal itself. Another form of authoritative websites, named FDA and CDA, are governmental public health agencies. Their purpose is to take an active role in issuing warnings and thwarting rumors as part of their regulatory functions. Their information tends to be the most reliable, but the article in [3] revealed that FDA might announce misleading information due to their limited experiments or not release a warning as early as it should be.

2.2 Collaborative Information Sharing

The user-generated content revolution has gained popularity through the wiki technology. Users can collaboratively edit and develop their content. Examples of a few well-known sites, such as *www.askdrwiki.com* and *www.ganfyd.org* are the sites that allow only physicians and medical experts to contribute to the sites. This is shown to be a reliable source for patients as well as the medical community at certain levels. Other forms of user-generated content where users can share health information are discussion forums. The knowledge in these sites depends considerably on user contributions. In the example of *www.taumed.com* and www.*medhelp.com*, participants answer questions or provide advice to one another. Other examples where patients express their opinion about their experiences of health care providers are *www.ratemds.com* and *www.healthgrades.com*. All mentioned sources share similar vulnerabilities. Frist, participants are physically anonymous to one another in sharing their content. There is not much participation in those sites. Therefore, the credibility of exiting content is doubtful. There are exiting mechanisms such as the reputation systems and peer monitoring to address such an issue

2.3 Social Network Sites

As social networks have gained popularity and become a part of the lives of people, the study [8] reported in May 2011 that there is a fair amount of health related social networking pages as follows: 1) 486 *YouTube* Channels related to health, 2) 777 *Facebook* pages, 3) 714 *Twitter* Accounts, 4) 469 *LinkedIn* social networks, 5) 723 *Four Square* venues, 6)120 *Blogs*. Furthermore, the specific *HSNS* have evolved to be an alternate solution for patients. *HSNS* are created for connecting patients to support one another. Patients could share their treatments, drugs and side effects. In the example of *www.patientslikeme.com*, members share their personal health information. In doing so, members can learn about their problem among one another including treatments and side effects. The issues of *HSNS* are quite similar to the issues in the collaborative information sharing. The difference is that users can obtain relatively more connections in the platforms. Hence, the accepted level of security mechanism is needed in such an application.

2.4 Multimedia

The multimedia sites are another source where patients obtain their information. The success of video sharing and the developing ubiquity of podcasts enable users to gather their health information. For instance, the study of [9] shows American hospitals have uploaded over 20,000 videos to *www.youtube.com*, or the sites like *www.icyou.com*. Similarly, the study also reveals that the issues of tags spamming and false information are presented in those sites.

For aforementioned services, a patient searching online for health information would not be able to easily distinguish a reliable review article from another that is biased or nonfactual. In such a scenario, the reliability of health information is crucial. Patients would like to know whether a claim or an article they find online is indeed trustworthy and which sources are more trustworthy than others. Based on our study, we focus on trustworthiness of health content so as to support patients in the decision-making process. Our study uses data from *www.epinion.com*, a user-generated content site where participants write reviews and rate several products based on their experiences.

CHAPTER 3. POSSIBLE ISSUES

3.1 Network Formation

The way to form connections of each *HSNS* requires several procedures. In some *HSNS*, users can easily obtain a large number of connections, while some require a lot of personal information to even become a member. In the case of *HSNS* that users easily obtain the connection, the connections tend to be weak ties, which implies that a user does not have much experience with such a connection. Malicious users can easily exploit such ties to manipulate their victims due to low cost compared to a strong tie.

3.2 Dissemination

Several *HSNS* have many different mechanisms that enable their participants to obtain desirable information. *Facebook*, for example, allows an individual to decide who else can view his or her information in his or her network, whereas in *Twitter* the information would be viewed by followers. The work of [10], researchers categorize the dissemination approaches into deterministic communication technique including distribution hierarchies such as in [11], [12], [13] and probabilistic communication techniques including epidemic based dissemination techniques such as probabilistic broadcast and flooding [14],

[15]. Each technique reflects how information flows from place to place. For a health scenario, spreading of false rumors may cause severe damage to many naive patients. Hence, dissemination approach in *HSNS* should be considered as another area where we should be concerned.

3.3 Standard Malicious Attacks

- Due to the nature of *SNSs* that allow individuals or organizations to create profiles for any purposes, malicious behaviors can exist in the systems; there are several classes of attacks which have been identified by the work of K. Hoffman [10] and can appear in the health scenario.
- Self-Promoting - Attackers manipulate their own reputation by falsely increasing it. For instance, drug companies may promote their products by hiring a group of people to write good reviews and ratings for their products.
- Self-Serving or Whitewashing - Attackers escape the consequence of abusing the system by using some system vulnerability to repair their reputation. Once they restore their reputation, the attackers can continue the malicious behavior.
- Slandering - Attackers manipulate the reputation of other nodes by reporting false data to lower their reputation.
- Denial of Service - Attackers may cause denial of service by either lowering the reputation of victim nodes so they cannot use the

system or by preventing the calculation and dissemination of reputation values.

CHAPTER 4. THEORETICAL BACKGROUNDS

4.1 Trust Metric Inspired by Measurement and Psychology

Measurement theory is a branch of applied mathematics that is useful in measurement and data analysis, including quantifying the difference between measured value and corresponding objective value. However, such a measurement may generally produce an error. Hence, a number of error approximation techniques have been introduced to represent the accuracy, precision or uncertainty of the measurement, including absolute error, relative error, confidence interval, and so on.

4.1.1 Psychology Implication

Trust is judgment made from people‘s impression toward others. The impression has been developed based on people‘s interaction and experience that their brain have repeatedly accumulated regarding other people. Such an impression assists humans to judge how trustworthy those people are. This formed trust can be used later in their decision making process. By the same token, physical measurements possess similar characteristics of human trust evaluation. However, the physical measurement can be improved its accuracy with many techniques, namely more precise equipment, different measurement methods, or repeating the measurement to reduce the error. This advantage inspired us to adapt the well-established and tested measurement theory in representing and computing trust relations in health social network applications.

4.1.2 Trust Metrics (Impression and Confidence)

m is introduced as a comprehensive summary of several *measurements* on a person's trustworthiness *say* Bob, which is evaluated by another person (say Alice). The evaluation is judged based on their real life experiences, including personal direct and indirect contacts in their social context, the concrete meaning of *m* depends on the specific scenario and application. For our health domain, we define *m* as a quality value (e.g. how good Bob is), a probability (e.g how likely Bob will tell the truth), and so on. However, the quality of *m* is similar to sampling in statistics in that the more incidents and experience Alice has on Bob, the more accurate *m* is, however, the accuracy must be depending to distribution of different impressions. A range of the distribution around the summarized trustworthiness measurement *m* can represent the best and worse judgment Alice had made on Bob. Such a range in fact refer how much Alice is confidence about her judgment on Bob, is similar to error in physical measurements, which represents the variance of the actual value from the summarized value. Therefore, *confidence*(*c*) is introduced. In psychology perspective, *c* represents how much a person is certain about his/her impression metric, while on statistical perspective, *c* determines how much away from real impression the measured one can be. Hence, we associate *c* with variance of measurement theory and statistics, in an inversely proportional manner. *c* is more easily to be assigned by people. However in order to utilize error propagation theory to compute transitive and aggregated trust (discussed in following sections), we must be able to convert confidence *c* to its error corresponding form. As a result, we further introduce another intermediate metric: range R, which is only used by the framework for computation. If we make *m* represent the measurement of trust, then R shows how much the expected best or worst trust can vary from the measured trust.

4.1.3 Value and Range of Trust Metrics

In trust metrics, we attempt to let users intuitively assign their impression regarding other users based on their own experience. We later employ *Likert-Scale* to convert the expression to a predefined value range of impression metric *m*, which is in the range 0 to 1 and so confidence do. As discussed in Section 4.1.2, the interpretations of their values can vary in many different circumstances. For our health scenario, we consider *c* as a percentage of known fact, whereas the percentage of uncertain fact would be $1-c$. Therefore, R should be the total impression range times the percentage of uncertain fact. Next we need to find the appropriate starting and ending value of R. For example, a trust of $m = 0.5$; $c = 0$ which represent the most neutral and uncertain trust, we would like the possible trust value ($m-r$ and $m+r$) could cover the whole range, i.e. the *real* impression value could be any number. On the other hand, if $c = 1$ which indicate highest confidence, the value of R would be zero which means both the worst and best expected impression equals to *m*. Following these guidelines, the relation between confidence and range can be simply defined as

$$R = 1 - c \qquad (1)$$

To better fit the error characteristic, radius *r*, which is half of range *R* is introduced. *r* shows how far the best or worst expected trust can be from the impression value *m*.

$$r = \frac{R}{2} \qquad (2)$$

Therefore, *m* is equivalent to measurement mean, and *r* is equivalent to square root of variance or standard error.

4.2 Trust Arithmetic Based on Error Propagation Theory

As discussed in 4.1.2, Alice is considered as a trustor who evaluates the trust level of Bob, whereas Bob is inversely called as trustee whose trust value have been evaluated by Alice. If Alice evaluate Bob and Bob also evaluate John, Indirect trust path is built by considering Bob as an intermediated node, and in reality a trustor can have more than one intermediated node. However, judgment of each node may present its error or uncertainty in statistics literature, which can be propagated and accumulated when system compute the trust value of a target trustee. In doing so, error propagation theory would come into the picture in order to summarize the overall error value of target trustee. In this section we would discuss the trust evaluation arithmetic based on error propagation theory using trust metric *m* and *c*, and how we adapt them to comply with psychological implications in our scenario. We will give an example of impression *m* computation equation, and how to generate corresponding confidence propagation equations. There are two basic types of trust prorogation operations: trust transitivity and trust aggregation.

4.2.1 Trust Transitivity

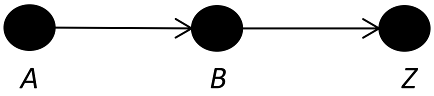


Figure 1 A Chain of Trust

We define Node *A* as the trustor node, and node *Z* as trustee target, and node *B* is an intermediate node which is considered as a gateway for trust information of target trustee. We define the operation of transitive trust as $\otimes$. Then node A's indirect evaluation of node Z via node B is represented as:

$$T_Z^{A:B} = T_Z^{A:B} \otimes T_Z^{A:C}$$

This can be viewed as a chain of trust path *A*-*B* and *B*-*Z* by using *B* as connecting from source to sink for trust transitivity. T_{AB} and T_{BZ} can be either direct trust or abstraction of transitive trust. Because our interpretation of trust metric: impression *m* and radius *r* correspond to the average and variance of a user's subjective evaluation based on past experiences, we apply the theory of error propagation for radius propagation after defining impression propagation equations. The equations for computing transitive trust should comply with psychological implications. Trust transitivity should obey the following properties, firstly $c_{ABZ} \leq c_{BZ}$. *A* cannot have more confidence than *B* just by taking *B*'s opinion. $m_{ABZ} \leq m_{BZ}$, Impression of z computed by the trust transitive should not bigger than viewpoint of *B* toward *Z*. without other supportive evidence, the impression would not get better than the original. The node which is closer to the trustor should have stronger influence on him. Hence, c_{AB} has more weight in c_{ABZ} than c_{BZ}.

Impression Transitive Equations: We define the indirect evaluation of node *Z*'s impression via node *B* that is computed as:

$$m_Z^{A:B} = m_B^A \ X \ m_Z^B \qquad (3)$$

Confidence Transitive Equations: Error propagation theory is adopted in this equation to compute the synthesized radius. The relative error of a production μ_1 μ_2 in statistics is computed as:

$$\left(\frac{\sigma_{production}}{\mu_1 \mu_2}\right)^2 = \left(\frac{\sigma_1}{\mu_1}\right)^2 + \left(\frac{\sigma_2}{\mu_2}\right)^2 + 2\left(\frac{\sigma_1}{\mu_1}\frac{\sigma_2}{\mu_2}\rho_{12}\right)$$

ρ_{12} is variance-covariance define the correlation between *m*1 and *m*2. When m_B^A and m_Z^B are independent, *A*'s opinion and *B*'s opinion are not correlated and ρ_{12} is equated to zero. We first start from computing absolute error:

$$\sigma_{production} = \mu_1 \mu_2 \sqrt{\left(\frac{\sigma_1}{\mu_1}\right)^2 + \left(\frac{\sigma_2}{\mu_2}\right)^2} \qquad (4)$$

Next we adapt this equation to our radius such that:

$$\sigma_{production} = m_B^A m_Z^B \sqrt{\left(\frac{r_B^A}{m_B^A}\right)^2 + \left(\frac{r_Z^B}{m_Z^B}\right)^2} \qquad (5)$$

Note that the relative error is applied as the argument being computed.

4.2.2 Trust Aggregation

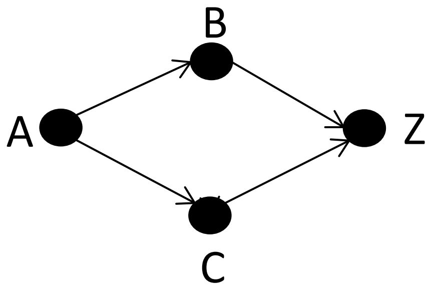


Figure 2 Trust Aggregation

Trust aggregation is introduced to summarize the propagated trust from multiple trust paths. We also use operator $\oplus$ to present trust operation aggregation. For instance, if two trust paths are presented to evaluate the trust score of node *Z*, the score of *A-B-Z* and *A-C-Z* would be aggregated for evaluation of node Z by computing as

$$T_Z^{A:Aggr} = T_Z^{A:B} \oplus \ T_Z^{A:C}$$

This aggregation is similar to combining two measurement populations together in statistics, in that their measurement mean could be an average based on population, and the variance would be the combination of two original variances. The main purpose of aggregation is to increase the confidence in decision-making process. Therefore, to rise and compromise the confidence, the opinions of each trust path is essentially deemed. Intuitively, if confidence is increased if similar opinion of information is presented from several paths, while it is worsened if different. Nevertheless, based on principle vulnerability may be introduced if a number of adversaries enhance their trust score by given similar opinions to target node. Confidence may drop if they provide contradicts opinions.

Based on Health information scenario, we must reply on the trust path with High confidence (high compensation of experiences). While aggregating, High confident path should not be highly suffered by trust path with low confidence. In other words, we give higher weigh on trust path with high confidence than low one.

Impression Aggregation Equation: When two indirect trust score are parallel, both of which give their opinions regarding to *Z*, for instance, node *B* and *C* both provide their direct score regarding node *Z* for node *A*. the impression could be computed as weighted average of paralleled impression(example shows for *A-B-Z* and *A-C-Z* paths) as following equation

$$\mathrm{m}_Z^{A:B} \otimes \mathrm{m}_Z^{A:C} = \frac{W_Z^{A:B} m_Z^{A:B} + W_Z^{A:C} m_Z^{A:C}}{W_Z^{A:B} + W_Z^{A:C}} \qquad (6)$$

W is the weight factor reflects the direct impression on intermediate node. We can define its value depends on scenario, for example, for our health decision making, we define $W = 1 / r^2$ which is identical to weighted mean. If there are limited amounts of sample, we can adjust the power of *r*. The trust path with higher confidence (low error) is favored. This is imitated from human behavior in that people tend to rely on other people with whom they have experiences.

Confidence Aggregation Equation: Our aim here is to apply measurement theory to capture decision making processes. If we aggregate multiple trust paths with weighted mean, the confidence will be increased comparable to single path. This is corresponding to the case that a user is certain about her judgment if she receive similar suggestions from multiple close friends regarding the same object.

$$\sigma \quad = \sqrt{\frac{\sum_{i=1}^{n} {W_{s_i}}^2 {\sigma_{s_i}}^2}{(\sum_{i=1}^{n} W_{s_i})^2}} \qquad (7)$$

Then if we replace *W* with *1 / σ*^{2}, we can get the formula (8), by which we can calculate in a recursive way.

$$\sigma \quad = \sqrt{\frac{1}{\sum_{i=1}^{n} {\sigma_{s_i}}^2}} \qquad (8)$$

Nevertheless, above equation does not capture the scenario that multiple highly trust nodes have different opinions regarding on the object. Hence, a conservative way is introduced to combine trust paths with dissimilar opinions. Here we represent trust path and its error as $m \pm \sigma$, which is an interval centered at *m*. We calculate combined *m* using arithmetic average and *σ* is chosen as the largest distance from centered point (combined *m*).

$$m = \frac{\sum_{i=1}^{N} m_i}{N} \qquad (9)$$

$$\sigma = \max\{|m - (m_i \pm \sigma)|\} \qquad (10)$$

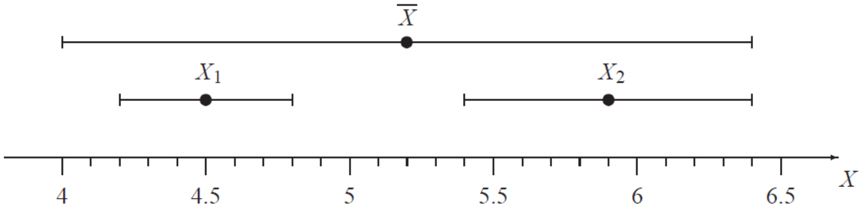


Figure 3 Conservative Way of Combination

The Figure 3 illustrates the foundation concept of Equation (10) that combines $X_1 \pm \sigma$ and $X_2 \pm \sigma$ in the conservative way. The combined mean covers all the range.

Confidence Aggregation Algorithm: Combination of multiple trust paths with their uncertainty requires us to utilize the Equation (8) (9) (10) into the algorithm in order to capture all decision making behavior as following procedures.

1) The aim of the first step is to filter an untrusted source out of the decision making process. We, therefore, consider *c* as a main factor whether a trust path is eliminated or not. We set certain score as a threshold and ignore a trust path that has less *c* score than the defined threshold. The threshold can be set depending on either a user or system administrator. The guideline for setting the threshold is based on scenario or a risk of information. For instance, a case of sensitive information, we must set high *c* as a threshold

2) The second step is to cluster the remain trust paths based on the similarity of *m*. the purpose of clustering is to maximize the confidence of

each group. The confidence will be much increased with the group that consists of many members, whereas not much increased with the group that consists of a few members.

There are several clustering techniques to apply here. Nonetheless, we simplify the solution by dividing trust paths into two groups which are [0, 0.5), [0.5, 1.0]. In each cluster, we assume that trust paths have similar *m*. then, we use Equation (6) and (8) to calculate *m* and *σ*. Consequently, we can obtain higher *c* than the threshold.

3) After obtaining *m* and *σ*, now each cluster has dissimilar *m*. Therefore, we treat both as different opinions and combine them together using Equation (9) and (10). Combination *m* will be on the middle of all groups, while the combination of *c* will be decreased due to conservative approach. Note that in certain cases, we may classify two closed *m* into two different groups, such as 0.49 and 0.5, but we can also get high confidence since the distance between them is small

CHAPTER 5. EXPERIMENTS AND ANALYSIS

We conducted several experiments to demonstrate how our trust framework applicable to health domain. Our study conducted on a real-world health social network dataset consists of five main tasks.

5.1 Data Crawling and Creating Social Networking

Validation of our framework is required to perform two main tasks: 1) we need to collect real data that represents how people interact in the health social network sites. 2) we present how we construct a trust network from the data. We elaborate the two tasks as follows:

First, we acquire health data by developing a crawler to retrieve the data from *www.epinion.com*. *Epinion* is the website where people come to share their experiences about several categories of products. The users' behavior of the site is describes as follows: Bob may have experiences about vitamin *A*, so he write a good review about it. Later, Alice come to the site and seeks the information about vitamin *A*. Next, she read Bob's review and rate Bob's review under a scale of 1-5. Since we pay interest on health domain, we narrowed down our data collection by crawling only rating and review of *wellness and beauty* categories, which consists of *Personal Care, Beauty Products, Hair care,* *Medicine Cabinet, and Nutrition Fitness* products. We started collected data in December 2011. In total, we extracted 3059 reviews. 788 out of them have been rated by other users, while there were 5081 users who rated other user's reviews.

Second, we construct the trust network by using the above collected data. Each user who either writes a review or rates a review represents a node in the network, while each rating denotes direct edge (direct trust) between nodes. For instance, Bob write a review about vitamin *A* and Alice rate Bob's review. The graph network is formed as follows: Alice node has a direct trust point out to Bob node. The direct trust between nodes has score of *m* and *c. m* present average of rating Alice give to Bob. *c* denote a number of rating Alice give to Bob. For this section, we obtain the trust network built from nodes and their relationship.

5.2 Verification of our Framework

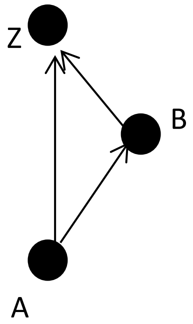


Figure 4 A pattern Retrieved for Verification

After collected the dataset, we verify the applicability of our trust framework based on the assumption that the *m* and *c* prediction result should be similar to the direct and *c* of real users. In this experiment, we compute the indirect *m* and the direct *m* expressed by real users, and compare them by Equation (11), and compute the similar process with c by Equation (12).

$$Diff(m) \ = \ |m_{direct} \text{-} \ m_{indirect}| \qquad (11)$$

$$Diff(C) \ = \ |c_{direct} \text{-} \ c_{indirect}| \qquad (12)$$

Later, we randomly selected pairs of users from the dataset when they have direct trust relation (review rating) and there is also a third user that can be used to form an indirect two-hop trust path the pair, so that we can compare the synthesized indirect trust with the original direct trust. Then, we synthesize the values based on our trust metric: *m* and *c*. For a trust relation from user *A* to user *Z*, impression *m* is assumed to be the average rating that *A* gives to *Z*'s reviews, and then converted to [0,1] range. Confidence *c* is synthesized from the number of review ratings given by *A* regarding *Z*'s, and is proportional to the square root of number of review rating (when number of review ratings increase, confidence tends to saturate) and topped at 1, shown as Equation (13).

$$\mathrm{C} \ = \ \min(1, \sqrt{1 * k \,\#\, (ReviewRating)}) \qquad (13)$$

According to our analysis of dataset, we define coefficient *k* equals 0.128, so 5 review rating will generate a confidence of 0.8. We are especially interested to compare direct (real) trust with indirect (synthesized) trust for cases of high confidence, which are the most important in any kind of decision making process. We use $Diff\,(m)$ to denote the differences between direct and indirect impression. By the same token, $Diff(c)$ denotes the difference between direct and indirect confidence. In Figure 5, it is shown the distribution of $Diff\,(m)$ that confirms that the indirect trust synthesized by our framework is a good approximation of direct trust expressed by real users.

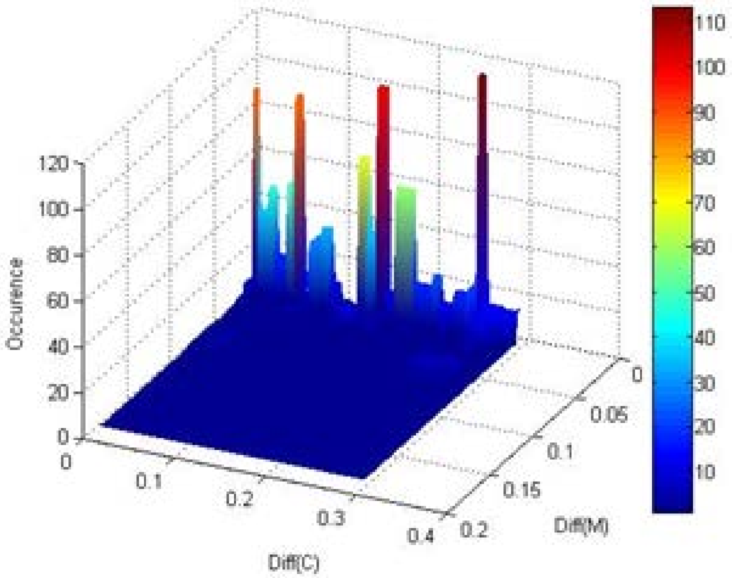


Figure 5 Difference between *m* and *c*

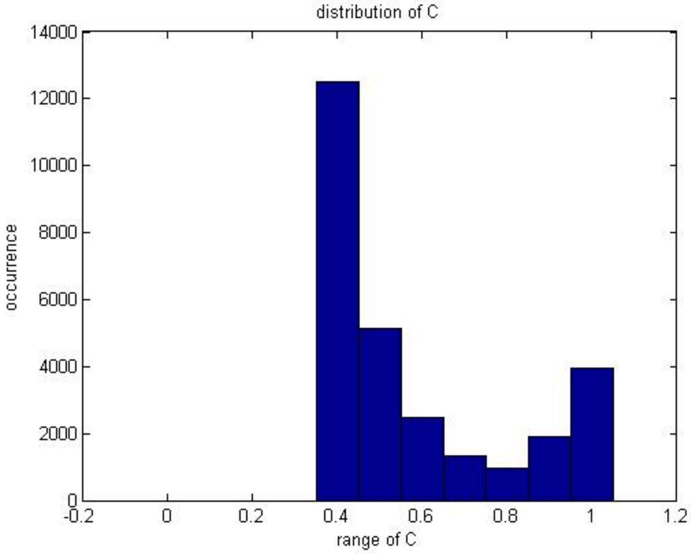


Figure 6 Distribution of Confidence without Aggregation

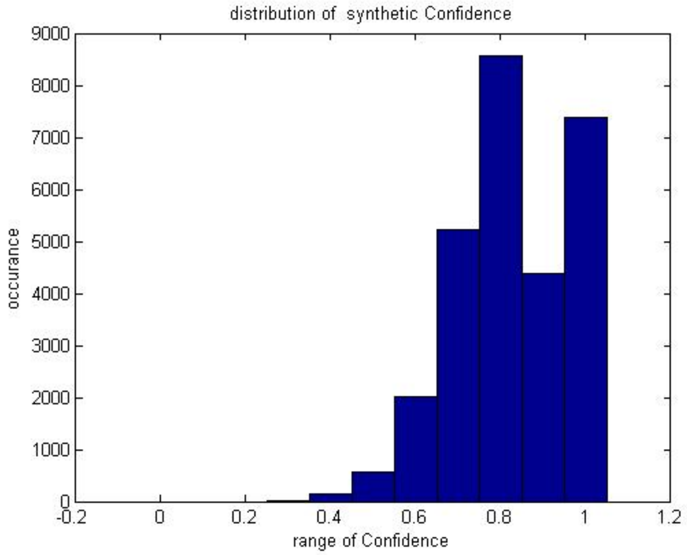


Figure 7 Distribution of Confidence with Aggregation

As we further explore the dataset, we are now interested in all aspect of our dataset (low confidence does not take into account). We found that confidence of each user is mostly low (in the range 0.4-0.5) as shown in Figure 6. This indicates that most trust paths are unlikely to reply on. This is challenging in decision making process. One possible method to address the issue, we aggregated the opinions of *A*'s neighbors giving to *Z* based on Chapter 4, Section 4.2.2. The result of confidence is improved as shown Figure 7.

We notice that not all the cases where confidence is raised. There is the case where confidence is compromised. This somehow does not help in decision making process. We therefore introduce a user who can raise the level of confidence, such as opinion leader or a user with high reputation In such a medical scenario, users' decision tends to intuitively count on a user whom they have much experience. We therefore select a group of nodes that has high reputation. We, in other words, describe them as a node with a plenty of rating made by other users, and assume that they are physicians. As a result, people will have a high level of confidence on them.

Each of expert nodes may have several levels of confidence, say in the range [0.7 - 1]. In decision making process, a node with the confidence of 0.7 does not have as much influence as does the node with 0.9. For certain cases, a node with low confidence may raise its confidence of its opinion by imitating a similar opinion from a node with higher confidence. Similarity, one expert (says a general doctor) may not assure to a specific issue beyond his or her expertise. S/he may ask to the expert to a given specific area. In doing so, s/he regain higher confidence to address the issue.

5.3 Attack Modeling and Consequential Effects

In this section, we investigate the detrimental effects of malicious behaviors on a network, such as Denial of Service and false rumors. For example, Company *A*'s aim is to promote their own product. As a result, it may hire dishonest users to rate or write a good review about its own products. On the other hand, it may hire attackers to sabotage competitors 'products so that many patients would resort to buying its company's products instead. In this case we assume that the attackers do not have many ratings to support their reputation. This experiment illustrates the possible impact that fake reviews can bring to the population. Based on Equation (14), if an attacker, say node *Z*, sent a message

(100) to several of their followers, the symbolic impact of that message received by a follower , say node *A* is computed as $100 * m_{Z}^{A:Syn} * c_{Z}^{A:Syn}$.

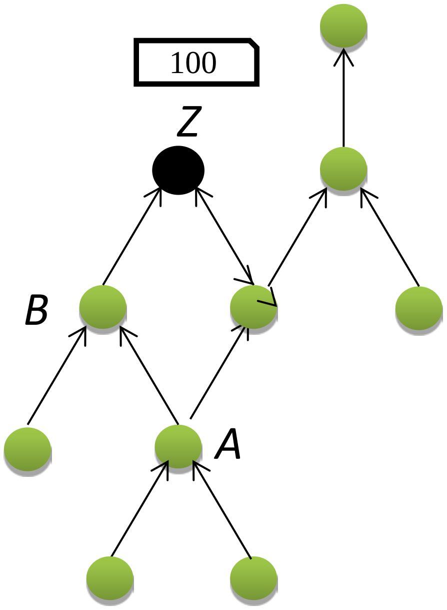


Figure 8 Illustration of How Node A Receives Message from Z

In the experiment, we sort users recorded in the extracted dataset according to the *Trust Power* they received, and divide them into three groups based on the range *Trust Power* they received.

- The first group of users is named *Power User* as they received plenty of *Trust Power* so they are generally known by many other users. This group of 50 Power Users is randomly selected from the pool ranging from the first to the 100th ranked.
- The second group of users is named *Moderate User* as they received some *Trust Power* so they are generally known by a few other users. This group of 50 Moderate Users is randomly selected from the pool ranging from the 300th to the 400th ranked.

- The third group of users are named *Less-known User* as they received very few *Trust Power* either because they are new to communities or they do not attract others' attention, but still have some history to some extent compared to entirely unknown. This group of 50 *Less-known Users* is randomly selected from the pool of ranging from the 600th to the 788th ranked.

Our simulation is calculating the total impact of each node if it has direct/indirect trust with one or many attackers. The simulation is set to compute three-hop maximum as indirect trust, and each node receives the impact computed by Equation (14).

$$Impact = \ InitialImpact \ * \ m \ * \ c. \qquad (14)$$

$InitialImpact$ is set to 100. The value of *m* and *c* is computed using one-to-one direct/indirect social relation between an attacker and a victim. For example, an attacker *Z* sends out misleading information, suppose user *B* is *Z*'s friend and user *A* is *B*'s friend, then user *B* and *A* will be victims. User *B* received an impact calculated by $100 * m_Z^B * c_Z^B$, whereas User *A* received an impact calculated by $100 * m_Z^{A:B} * c_Z^{A:B}$. Additionally, if there are multiple indirect trusts, the value of *m* and *c* need to be computed by the aggregation model (e.g $m_Z^{A:Syn} * c_Z^{A:Syn}$). Since there are 50 malicious nodes, the range of damage one user can receive is from 0 to 5000. Figure 9 illustrates overall impact of all nodes in each type versus the number of attackers. Clearly, the overall impact increases at the beginning then becomes saturated when the number of attackers increases. This pattern corresponds to the typical human behavior that when information sources are limited, we tend to consider every one of them. Then, when a large number of sources converge, our mind generally would not change much because of the input from a few new sources. Figure 10 and 11 demonstrate the characteristic of the trust framework that has a good defense against Denial of Service. We assume that *Less-known Users* have a high possibility to be attackers due to easiness of generation. Hence, if we use *c* value as a threshold, the sink nodes for each victim‘s view that have lower direct /indirect *c* than the threshold will be filtered out. Figure 11 illustrates the impact of Less-known Users are decreased after applying several values of threshold, but it still has a high trust path. The result implies that *Less-known Users* are less likely to be attackers. In Figure 12, we inject a group of nodes that has low *c* (less than 0.3) in the network. This group has an impact to a certain level. We consider this group as attackers. However, after applying 0.3 threshold, the group is filtering out.

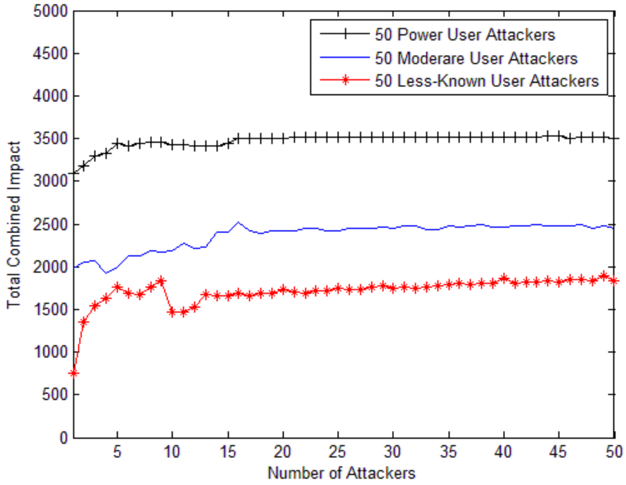


Figure 9 Total Impact of Attackers on Epinions

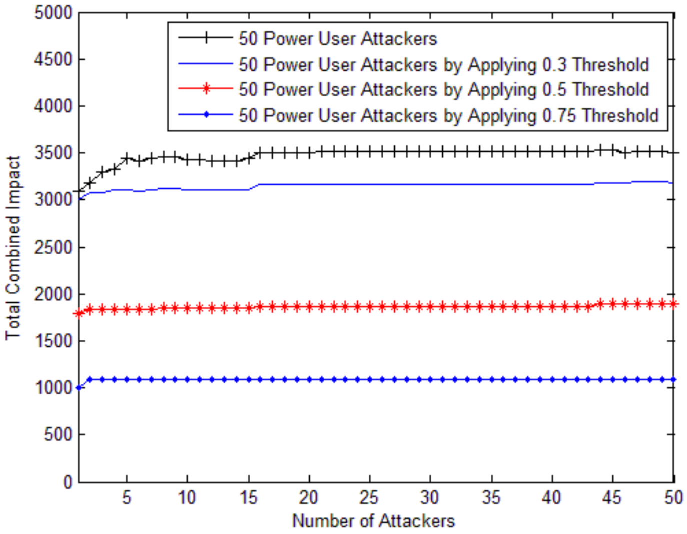


Figure 10 Total Impact of Power User Attacker by Applying Thresholds on Epinions

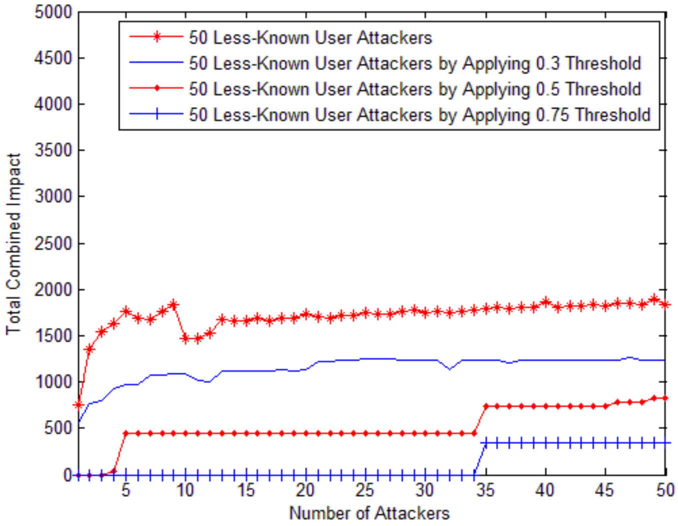


Figure 11 Total Impact of Less Known User Attackers by Applying Thresholds on Epinions

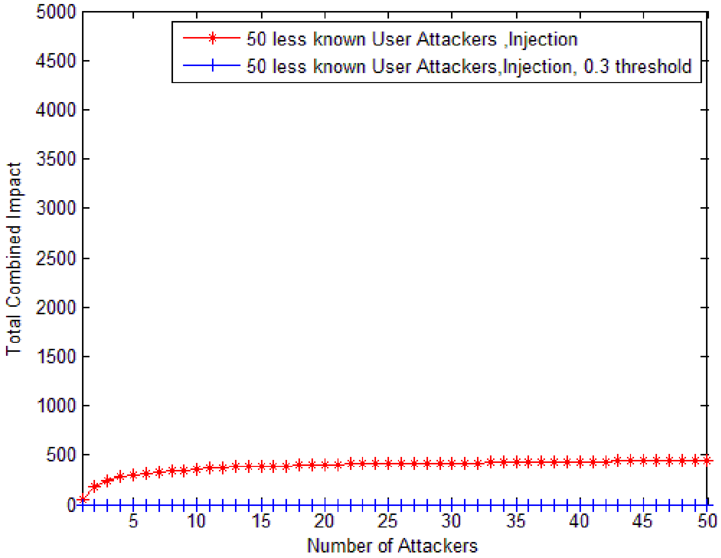
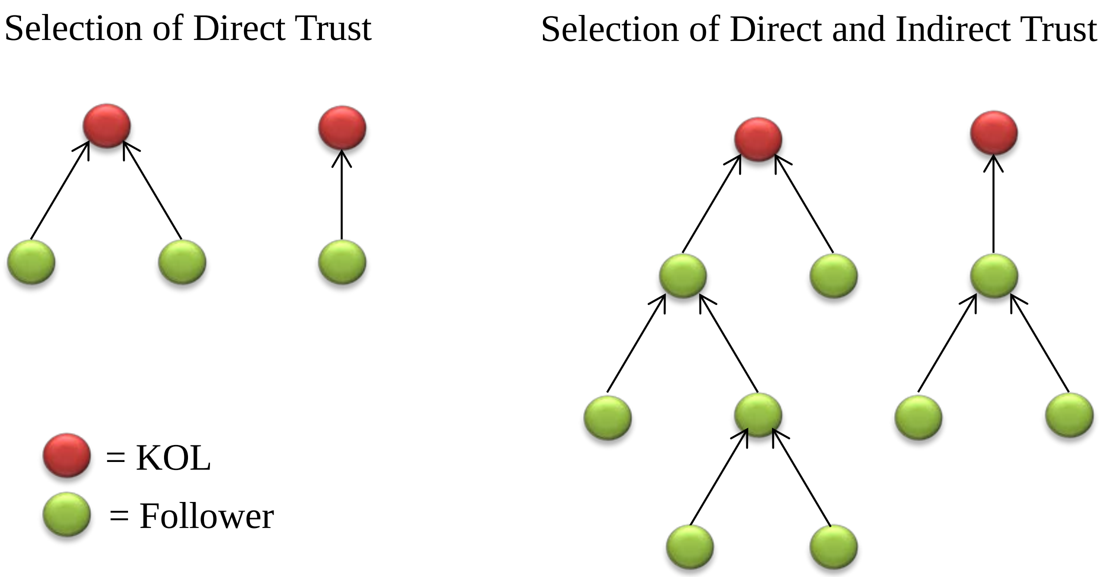


Figure 12 Total Impacts of Fake User Attackers

5.4 Pharma Marketing Model

As physicians decide which drugs to prescribe for their patients on a daily basis, the decision probably has the largest influence on medical industry revenues. As a result, many healthcare advertising companies have several tools to track physicians’ prescription patterns. Regarding these patterns, they will rank the influence power of each physician toward patients into several groups. The physicians of the highest ranking group are considered as Knowledge Opinion Leaders (*KOLs*). The companies can exploit such data by hiring *x* advertiser-*KOLs* so that the overall effect of their advertisement can be improved, or maximized.

In these experiments, we show how an advertiser could improve the advertisement effect on health consumers based on two solutions. A first solution is to select advertiser-*KOLs* according to the number of their received review ratings (i.e. only direct trust pointing to a user). However, the vulnerability of this scheme is easily identified as potential cheating users intentionally generate fake IDs by using Botnets to promote a given user. The second solution is to select advertiser-*KOLs* by considering trust relation (i.e. direct and indirect trust pointing to a *KOL*). This solution utilizes all possible network information.


Selection of Direct Trust

Selection of Direct and Indirect Trust

Figure 13 Difference between Two Selection Methods

To compare these two approaches, we first sort all users based on two different criteria: 1) according to the number of received review ratings, and 2) according to the total trust (calculated by our trust framework) by which each user can affect the population. Furthermore, we select users in two different criteria exclusively. These different criteria indicate that a selected node appearing in one criteria will not appear in the other. We also consider three hops as the maximum level of indirect trust to compute the effect. Last, we consider two possible types of advertisement effects (*AD effect*): *simple* and *intelligent* as measurements.

The simple *AD effect* consists of selecting advertiser-KOLs (*ADer-KOLs*) who simply send the same slogan or message to the network. A user who receives the message from multiple *ADer-KOLs* will not get a *combined AD effect* exceeding the value received from highly trusted advertiser-users. However, the *combined AD effect* would be reinforced if received from multiple highly trusted sources. For instance, if user *A* receives an advertisement from users *B* and *C*, and if user *A* has high confidence in both *B* and *C,* two possible outcomes are presented. 1) *B* and *C* have similar impressions (i.e. the difference between $m_{product}^{A:B}$ and $m_{product}^{A:C}$ are small) and the *combined AD effect* will be reinforced. 2) They have contradicting impressions and the *combined AD effect* will be compromised.

The intelligent *AD effect* consists of selecting *ADer-KOLs* who express their own impressions and describe various aspects of the products in a personalized manner. For instance, Each *ADer-KOL* shares an opinion about a particular drug. Hence, we assume that messages are independent, and for simplicity, the *combined AD effect* for each *ADer-KOL* will be the sum of each user receiving *AD effects*.

The results of *simple* and *intelligent AD effects* are shown in Figure 14 and Figure 15 respectively. Both results illustrate the power of our trust framework in the improvement of advertisement on health social networks.

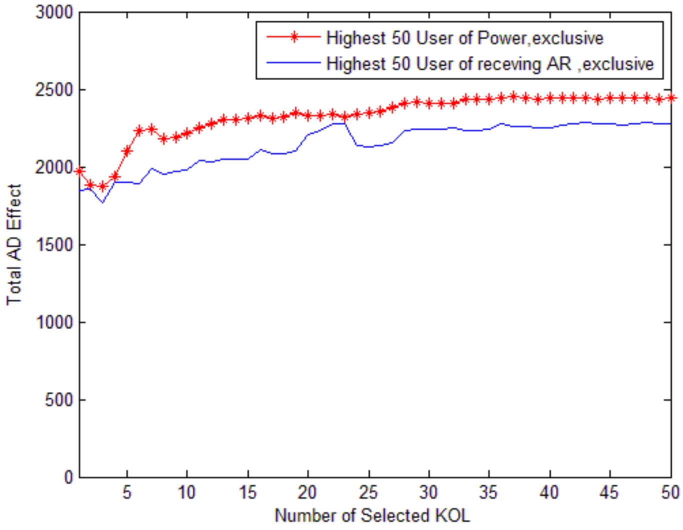


Figure 14 Simple AD Effect

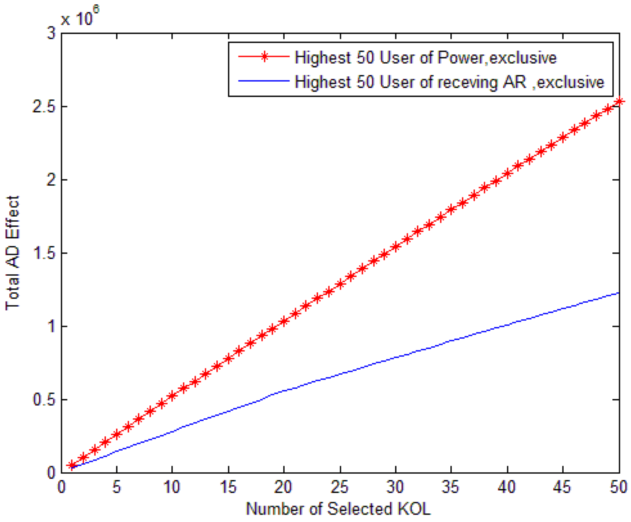


Figure 15 Intelligent AD Effect

5.5 Contradiction of Knowledge Opinion Leader (*KOL*)

In HSNS, *KOL* provides recommendations to their patients (followers). In reality, patients can express different levels of trust in any physicians. For instance, patient *A* may have more confidence in Doctor *B* than Doctor *C*. In the decision making process, patient A would rely on Doctor *B* instead of Doctor C.

Nonetheless, in certain cases, a patient may have multiple highly trusted sources who give different opinions. Because of this, the confidence of the patient is decreased. In this section, we focus our study on the impact of contradictory opinions among *KOL* toward their followers in two scenarios.

The first experiment illustrates the impact of followers if a group of *KOL* presents different opinions. We sort users recorded in the extracted dataset according to the *Trust Power* (the same process as Section 5.4) they received. Later, we select the top 10 users as *KOL* (doctor) and set each doctor to send recommendations to their followers. The recommendations consist of positive 100 and negative 100. Each doctor will send either one of the recommendations. This refers to contradictory opinions. An impact to each follower is computed based on Equation (14). For example, Doctor *Z* send -100 to patient *B*, who has direct trust with Doctor *Z*. Hence, the impact to patient B is $-100*\mathrm{m}_{\mathrm{Z}}^{\mathrm{B}:} * \mathrm{c}_{\mathrm{Z}}^{\mathrm{B}:}$.The message (-100 or +100) is analogous to the situation when a doctor suggests that a patient do something. +100 can be interpreted as a doctor tells a patient to get the treatment *A* for his or her disease. -100, on the contrary, indicates the doctor told the patient to get the treatment B for his or her disease. The *Trust Power* of Doctor *C* is computed based on the sum of each patient's impact, which is similar to the *intelligent AD effect* scenario, whereas total combined impact is calculated from the sum of power of all doctors in the network.

The experiment captures the shifting of opinion among 10 doctors. We present 10 stages of 10 doctors giving recommendations (+100 or -100). The 1th stage consists of the first doctor giving +100, while the other nine doctors give -

100. The 2th stage consists of the first two doctors giving +100, whereas the other eight doctors give -100. The 3th stage consists of the first three doctors giving +100, while the other seven doctors give -100. We resume with this process until the last stage, consisting of all ten doctors providing the same opinion +100. In Figure 16, the total combined impact is increasing in each stage of 10 doctors' opinions. The total combined impact obviously is shifting from negative to positive. On the other hand, Figure 17 shows that the number of followers who receive positive opinion (+100, 0) are increasing in each stage. Interestingly, after the 4th stage, the increase presents significant shifting. And after the 6th stage, all connected nodes receive all positive opinions. For this dataset, we conclude that 60 percent of the doctor group bring consensus among their followers. Alternatively, we chose the number of reviews as a method to select *KOL*. In Figures 16 and 17, the blue graph illustrates the impact of number of reviews as a method of selection. The impact would not be as good as *Trust Power.*

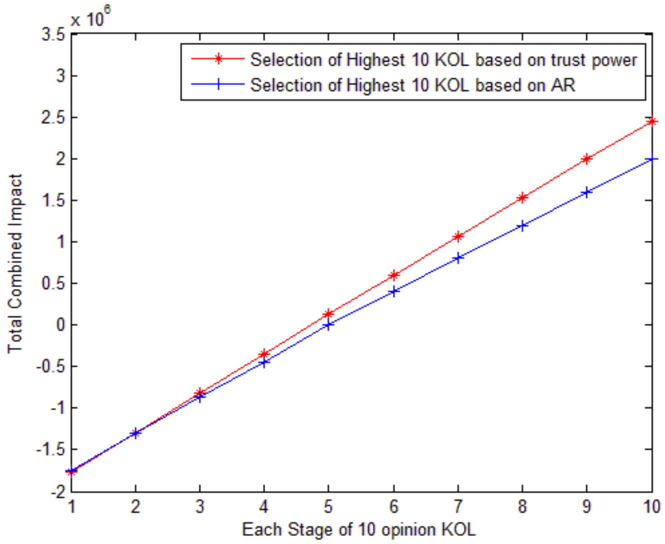


Figure 16 Combined Impact for 10 *KOLs*

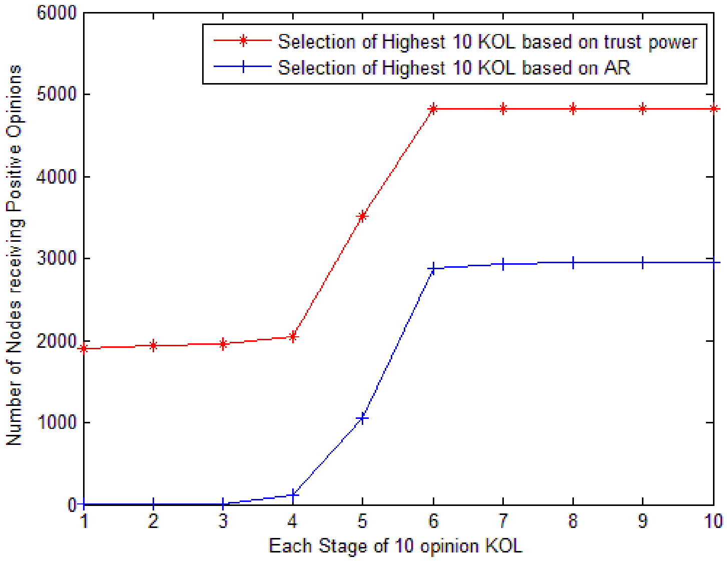


Figure 17 Number of Nodes Receiving Negative Opinions

In the second experiment, we performed a similar process. It is to compute both the impact and number of nodes that receive a positive opinion, but we would like to view them in another perspective. This is to capture the role of each node in an introduction of new medicine. The role in this case refers to

Trust Power score. A node with high *Trust Power* clearly makes higher influence compared to the lower one. We compare this experiment to a scenario where certain nodes propose new health information. Whether the information will make an impact or not depends upon many factors, one of which is *Trust Power* score. The experiment illustrates total impact when each node with a different degree of *Trust Power* sends a message (-100 or +100). In Figure 18, when the 1, 2th nodes (*Less- Known Users*) send +100, the impact through the network is presented in certain levels. When the 3th node (*Moderate User*) sends -100, the impact is dropping. The impact increases again when the 4th node (*Moderate User*) sends +100. Interestingly, when the 6th node that receives the highest *Trust Power* sends +100, the impact is significantly increased and becomes higher and higher when the 7, 8, 9 and 10th nodes that are considered as the top *KOL* send +100. Figure 19 illustrates the same process, but instead of combined impact, we present a number of nodes that receive positive nodes. Obviously, after the 6th node sends +100 message, the rest of the network turn to follow the positive opinion. Figures 20 and 21 have a similar process as 18 and 19 respectively. One difference is the 6th to the 10th nodes are not selected based on *Trust Power*. We select those nodes from the number of reviews instead. This type of selection is vulnerable to attackers. This does not have an effect on the network.

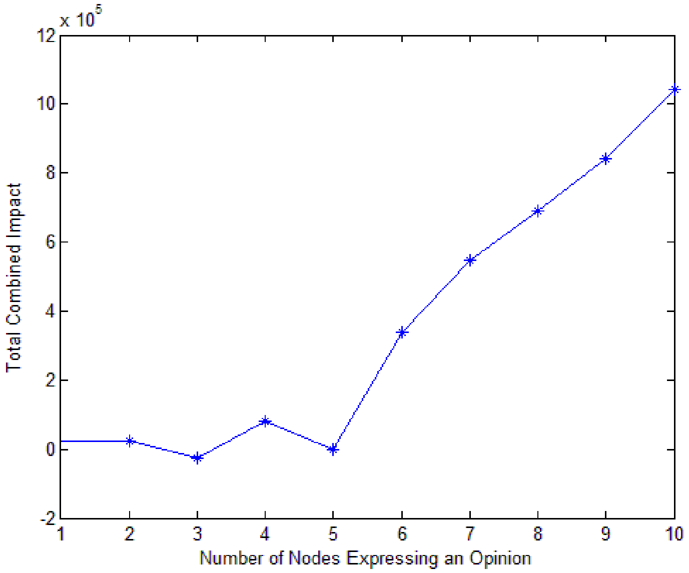


Figure 18 Impact of Contradictory Opinions

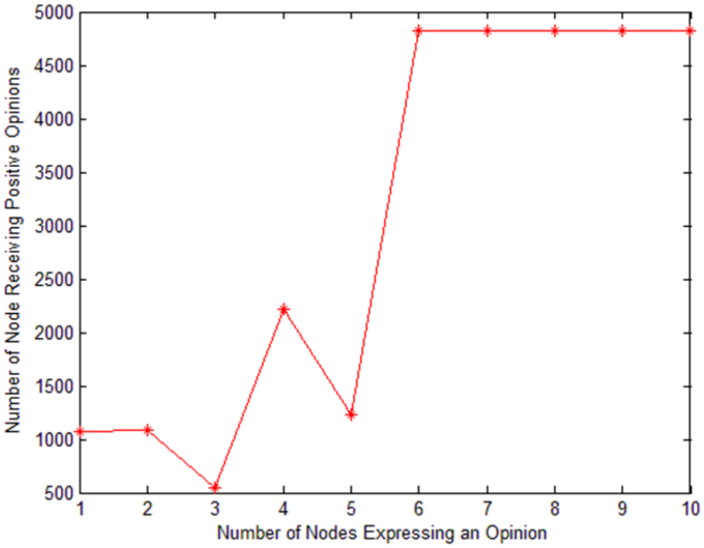


Figure 19 Number of Positive Nodes toward Conflict Opinions

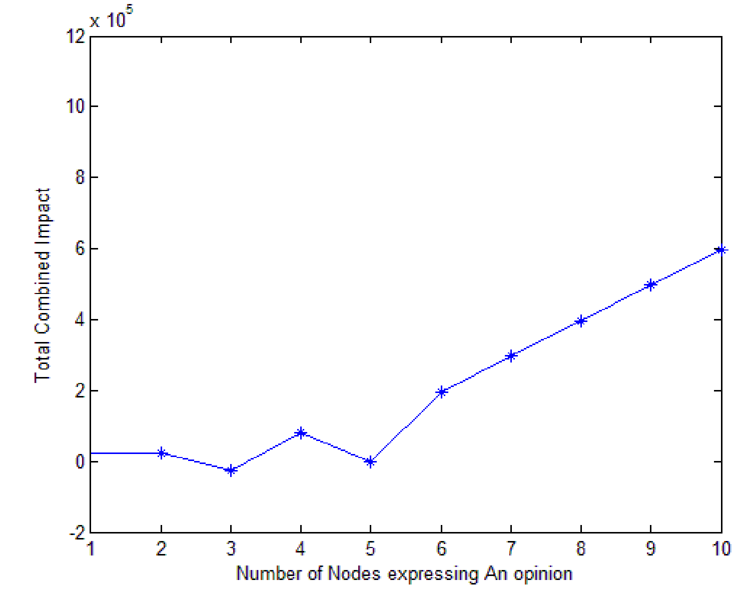


Figure 20 Impact of Contradictory Opinions with Fake Nodes

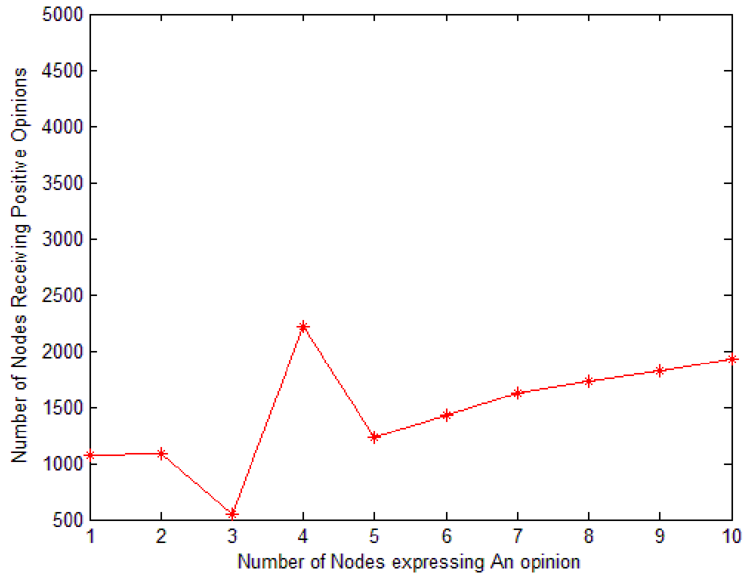


Figure 21 Number of Positive Nodes toward Conflict Opinions with Fake Nodes

CHAPTER 6. COMPARISION TO PREVIOUS WORKS

More than a thousand researchers have been working in the area of trust management. Most of them built their frameworks on the subjective direction. Our framework, on the other hand, uses the measurement theory which has been proved and accepted for more than a century. In this section, we compare our framework with another work in two aspects: robustness and identification of influencers for marketing tools.

6.1 Robustness to Attackers

Robustness of our framework is compared with the work of Fullam et al. [28]. Their work introduces a framework that justifies the reliability and uncertainty of information sources. Their direction is similar to our framework in that it consists of the trust score and its uncertainty. Thus, it is suitable for us to compare the performance. We used the same experiment as in Section 5.3 to capture the impact of attackers. Figure 22 illustrates that after the presence of the 20th node, the impact of our framework becomes saturated. This resembles human behavior in that when many sources of information are presented, a decision maker relies on a few sources that s/he trusts the most, and neglects the rest of the information. On the contrary, the other framework presented the impact as still in a stage of increasing. We consider the framework as vulnerable. *Less-known Users* (a high tendency for malicious nodes) can increase the trust score of information sources to manipulate the rest of the nodes.

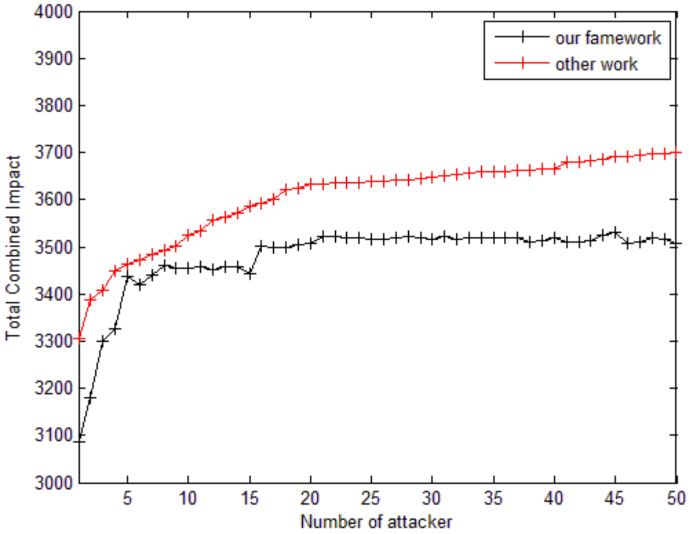


Figure 22 Comparison of Robustness with a Previous Work

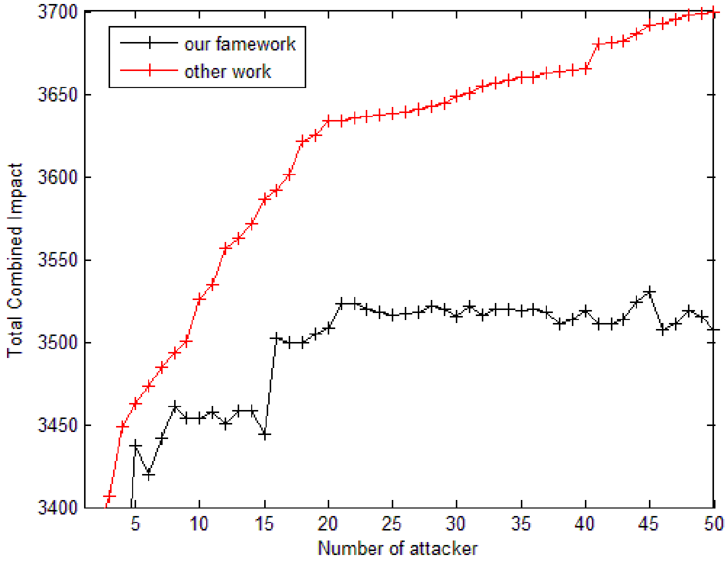


Figure 23 Zooming Comparison of Robustness

6.2 Identification of Influencers

We compare two approaches of a selection by again using combined impact as measurement. The approach that has higher impact is the better approach. We compare *Trust Power* selection with *In-degree* selection. Figure 24 illustrates that the former outperforms the latter. In addition, if we inject nodes that have the high score of *in-degree*, but low *Trust Power* score, the result of the impact is not as great as the *Trust Power* selection as shown in Figure 25. This implies that the *In-degree* approach also is vulnerable to attackers (*Less-known Users*).

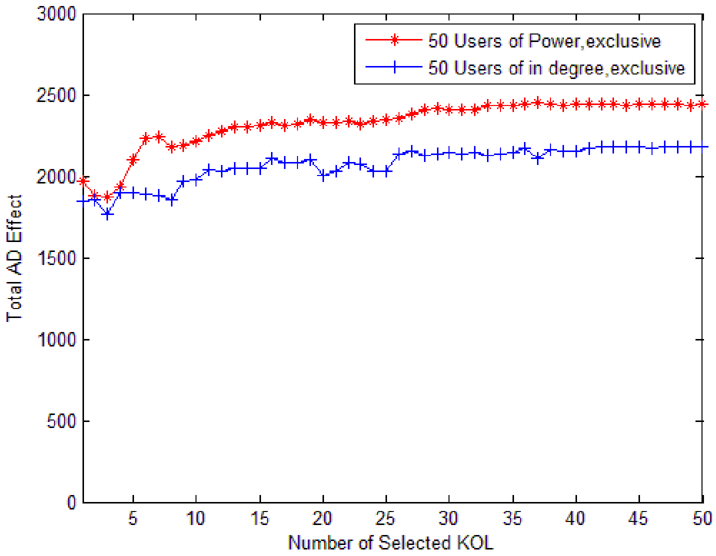


Figure 24 Comparison of Selection Methods

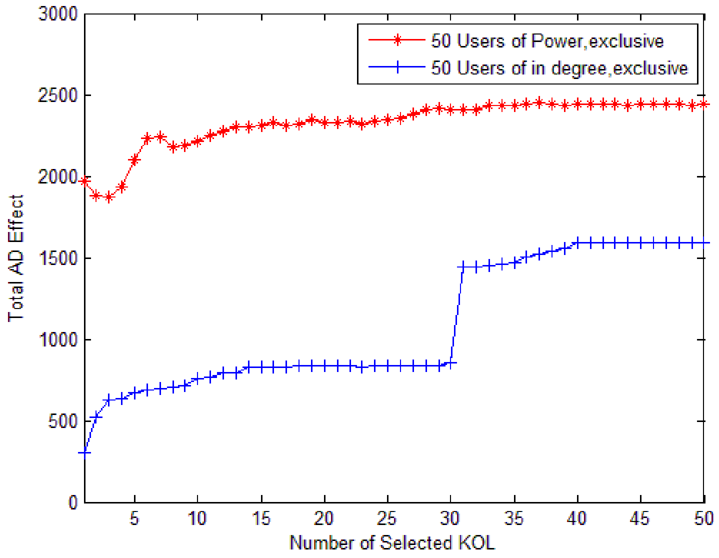


Figure 25 Comparison of Selection Methods with Fake Nodes

CHAPTER 7. RELATED WORKS

There are several aspects we should consider in an assessment of credibility of information in *HSNS*. In this section, we identify as follows.

7.1 The Trustworthiness of Source and Claim

Weitzel et al. in [16] presents a framework that measures a trustworthiness of health information source. Their study was conducted based on health websites which also appear in *Twister* SNS. The trust model is basically built behind the hypothesis of two equations, namely reputation factor and a set of quality indicators. The former equation is built from harmonic mean which is extended from their previous work [17] that studied topological structure of *Retweet* weighted ties. On the contrary, the latter equation is created from arithmetic mean by considering several technical criteria in medical domain, such as those interactivity and certification in the sites as a parameter. At the end, the framework applied harmonic means to aggregate the results of two equations behind the notion that mean does not reflect the quantity desired

HealthTrust developed by Fernandez-Luque et al. [9] is introduced to analyze the reliability of information in the diabetes online community. The core algorithm is created from Hyperlink-Induced Topic Search (*HITS*) for ranking the most authoritative source. The correctness of the algorithm begins with extracting health information from *YouTube* channels, particularly the diabetes online community. Later, they utilize *HealthTrust* to rank the most authoritative diabetes channels. The ranked list of channels from *HealthTrust* was compared with the list of the most relevant diabetes channels from *YouTube*. Two healthcare professionals are selected to classify the channels based on whether they would recommend the channel to a patient. The result of human would be considered as a benchmark in performance of two algorithms. Based on precision measurement, *HealthTrust* performed several times better than *YouTube* for filtering out the worst channels. However, a limitation of their study is very limited data sets for evaluation which make their results weak from a statistical point of view.

Vydiswaran et al. [18] propose another feasible method to predict the trustworthiness of a medical claim based on experiences shared by users in health forums and mailing lists. Their objective is to address the question whether community-generated text can be reliably used to predict reliability of the claim. The claim scores can be used to rank related claims on their relative trustworthiness. They further extend the notion of trustworthiness to a site (or equivalently, a database of claims from the site) and propose a scheme to rank sites based on aggregating the trust scores of claims from the site. The experiments demonstrated that community knowledge can be utilized to help users distinguish reliable medical claims from unreliable ones.

The study presented by He et al. [19] mainly explores the security and trust system related to the area of wireless medical sensor networks. They identify the security and performance challenges facing a sensor network for wireless medical monitoring. They introduce an attack-resistant and lightweight trust management scheme so called *Retrust*. The scheme is basically computed based on the beta-function-based method, but they modify some parameters which are similar to the work of Feng et al. [20]. The model is resilient to attack.

Clifford et al. [21] focus their study on trustworthiness in pervasive medical applications. In the papers, they described how to apply their previous works, so called *Solar Trust Model* [22][23] to the health scenario. The model is basically built to determine the relative trustworthiness of data from many potential sources based on the assumption that users may join and leave a network randomly.

Alhaqbani et al. [24] propose the model to determine the trustworthiness of information in an Electronic Health Record system named Time-variant medical Data Trustworthiness. The model evaluates the trust score of an agent based on both direct experiences that are computed from its own record retrieved from its database and external sources which are retrieved from neighbor’s experiences and healthcare reputation center.

Levy et al. [25] developed a prototype of healthcare social network system (*the Husky eHealth2.0*) to enhance the system's privacy control scheme. The system considers several important factors related to the privacy requirements in e-Health 2.0 applications, including user availability, user popularity, user participation, and user level of competency. They also developed and implemented a trust-aware tag-based privacy control scheme based on these factors. They evaluated their prototype via online survey.

7.2 Finding and Monitoring Influential Users

Yang et al. in [26] focus their study to understand how information is spread in health online communities. The aim of their research is to understand how the public reacts toward epidemics. In doing so, they propose a framework to monitor and identify influential users from online healthcare forums. They developed a mechanism to identify and construct social networks from the discussion board of an online healthcare forum. They invent an algorithm so called *UserRank* which results from combination of link analysis and content analysis techniques so as to identify Influential users. They evaluate the quality of their algorithm based on precision and rank distance which require human as standard ranking. Their experimental results show that the technique outperforms *PageRank*, *In-degree* and *Out-degree centrality* in identifying an influential user from an online healthcare forum.

M. Paul et al. [27] introduced Ailment Topic Aspect Model Plus (*ATAM+*) which analyzes Twitter messages about influenza tracking results. *ATAM+* model is improved by using prior knowledge, and reports results for several new applications which are geographic syndrome surveillance for multiple ailments (tracking illness over time and location), correlating behavioral risk factors with ailments, and analyzing correlations of symptoms and treatments with ailments.

CHAPTER 8. CONCLUSION AND FUTURE WORK

As *HSNS* have been a crucial tool for patients to consume health information, harm caused by false information can cause severe damage to health. In this thesis, we present how to apply our previous trust framework to assist individuals to filter unreliable information, to help system administers to improve their advertisement tool and to model a consequential effect in contradict opinion of *KOL*. In The future research, we would like to investigate the data from twitter or Facebook, and particularly answer the question what factors do patients use regarding health decision? and do physicians influence patients differently from non-physicians in *HSNS*?

REFERENCES

REFERENCES

[1] Pew Research Center. (2011) The Online Health Care Revolution: How the Web Helps Americans Take Better Care of Themselves, Retrieved from : http://www.pewinternet.org

[2] Survey by Harris Interactive. (2009) Patient Choice an Increasingly Important Factor in the Age of the Healthcare Consume, Retrieved from : http://www.harrisinteractive.com/Insights/HarrisVault.aspx

[3] Walker, E. (2010) Drugmaker FDA, Both Admit to Mistakes in Recall, Retrieved from : http://www.medpagetoday.com/ProductAlert/OTC/22504

[4] Theng, Y., Goh, L., Tin, M., Sopra, L., and Kumer, L. (2012) Trust Cues Fostering Initial Consumers Trust: Usability Inspection of Nutrition and Healthcare Websites, International Health Informatics Symposium, Proceedings of the 2^{nd} ACM SIGHIT, doi : 10.1145/351092.351099

[5] Atwood, K. (2009) An Overview of Misleading Health Information Found on WebMD, Retrieved from : http://getbetterhealth.com/an-overview-of-misleading-health-information-found-on-webmd/2009.08.27

[6] Gallegos, A. (2011) Doctors Legal Remedies Can Defect Online Attacks, Retrieved from : http://www.amaassn.org/amednews/2011/12/12/prsa1212.htm

[7] Zhang, P., Durresi, A. (2012) Trust Management Framework for Social Networks, ICC-2012, in Proceedings of the IEEE International Conference on *Communications*, Retrieved from : http://cs.iupui.edu/~durresi/1papers.html

[8] Bennett, E. (2011) Hospital Social Network List, Retrieved from : http://ebennett.org/hsnl/

[9] Fernandez-Luque, L., Karlsen, R. and Melton, G. (2011) HealthTrust: Trust-based Retrieval of YouTube's Diabetes Channels, Conference on Information and Knowledge Management, doi : 10.2196/jmir.2237

[10] Josang, A., Ismail, R. and Boyd, C. (2005) A Survey of Trust and Reputation Systems for Online Service Provision, Decision Support Systems, vol. 43, no. 2, pp. 618-644, doi : 10.1016/j.dss.2005.05.019

[11] Walsh, K. (2006) Experience with An Object Reputation System for Peer-To-Peer File Sharing, Networked System Design and Implementation, doi : 10.1109/TKDE.2004.1318566

[12] Williams, J., Weber-Jahnke, J. (2010) Social Networks for Health Care: Addressing Regulatory Gaps with Privacy-by-Design, 8th annual international conference on privacy, security and trust, doi : 10.1109/PST.2010.5593252

[13] Dimitriou, T., Karame, G. and Christou, I. (2007) SuperTrust: A Secure and Efficient Framework for Handling Trust in Super Network, ACM Symposium Principles of Distributed Computing, doi : 10.1016/j.dss.2005.05.019

[14] Kamvar, D., Schlosser, T. and Garcia-Molina, H. (2003) The EigenTrust Algorithm for Reputation Management in P2P Networks, the 12th international conference on World Wide Web, doi : 10.1145/775152.775242

[15] Zhou, R., Hwang, K. (2007) PowerTrust: A Robust and Scalable Reputation System for Trusted Peer-to-Peer Computing, Parallel and Distributed Systems, IEEE Transactions on Information Technology, vol. 18, no. 4, pp. 460-473, doi : 10.1.1.178.4335

[16] Weitzel, L., Quaresma, P. and Oliveira, J. (2012) Evaluating Quality of Health Information Sources, IEEE International Conference on Advanced Information Networking and Application, doi : 10.1109/AINA.2012.41

[17] Weitzel, L., Quaresma, P. and Oliveira, J. (2011) Measuring Node Importance: A Multi-criteria Approach, in WWW/INTERNET, doi : 10.1109/GreenCom-CPSCom.2010.26

[18] Vydiswaran, V., Zhai, C. and Roth, D. (2011) Gauging the Internet Doctor: Ranking Medical Claims Based on Community Knowledge, Data Mining for Medicine and HealthCare, pp. 42-51, doi : 10.1145/2023582.2023589

[19] He, D., Chen, D., Chan, S., Bu, J. (2011) ReTrust: Attack-resistant and Lightweight Trust Management for Medical Sensor Networks, IEEE Transactions on Information Technology in Biomedicine, vol. 16, no. 4, doi : 10.1109/TITB.2012.2194788

[20] Feng, Q., Liu, L., Dai, Y. (2012) Vulnerabilities and Countermeasures in Context-Aware Social Rating Services, ACM Transactions on Internet Technology, vol. 11, no. 3, doi : 10.1145/2078316.2078319

[21] Clifford, M., Bishop, M. (2011) Trust of Medical Device, Application, and Users in Pervasive Healthcare, PErvasive Technologies Related to Assistive Environments, no 54, doi : 10.1145/2141622.2141686

[22] Clifford, M., Lavine, C. and Bishop, M. (1998) The Solar Trust Model: Authentication without Limitation, Computer Security Applications Conference, Proceedings of the 14th Annual Conference, pp. 300-307, doi : 10.1109/CSAC.1998.738650

[23] Clifford, M. (2002) Networking in The Solar Trust Model: Determining Optimal Trust Paths in a Decentralized Trust Network, Computer Security Applications Conference, pp. 271-28, doi : 10.1109/CSAC.2002.1176298

[24] Alhasbani, B., Fidge, C. (2009) A Time-Variant Medical Data Trustworthiness Assessment Model, e-Health Networking, Applications and Services, Proceedings of the 11th International Conference, pp. 130-137, doi : 10.1109/HEALTH.2009.5406198

[25] Levy, K., Sargent, B., Ba, Y. (2011) A Trust-Aware Tag-Based Privacy Control for eHealth 2.0, Home for 12th Annual conference on IT Education, Retrieved from : http://sigite2011.sigite.org/?presentation=a-trust-aware-tag-based-privacy-control-for-ehealth-2-0

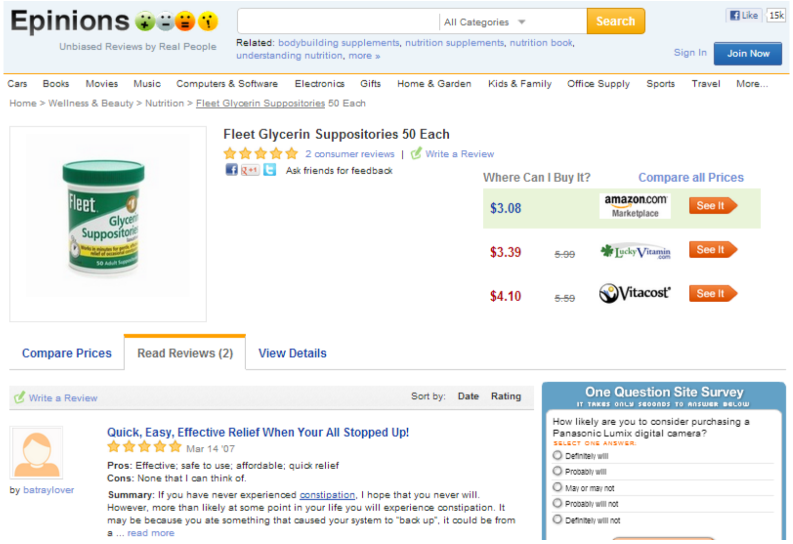
[26] Yang, C., Tang, X. (2010) Identifying Influential Users in an Online Healthcare Social Network, Intelligence and Security Informatics, pp. 43-48, doi : 10.1109/ISI.2010.5484779

[27] Paul, M., Dredze, M. (2011) You Are What You Tweet: Analyzing Twitter for Public Health, International Conference on Weblogs and Social Media, Retrieved from : http://www.cs.jhu.edu/~mdredze/publications/twitter_health_icwsm_11.pdf

[28] Fullam, K., and Barber, K. (2004) Using Policies for information Valuation to Justify Beliefs, Autonomous Agents and Multiagent Systems, pp. 404-41, doi : 10.1109/ISI.2010.5484779

APPENDIX

APPENDIX



Related: bodybuilding supplements, nutrition supplements, nutrition book, understanding nutrition, more »

Home > Wellness & Beauty > Nutrition > Fleet Glycerin Suppositories 50 Each

Fleet Glycerin Suppositories 50 Each

2 consumer reviews | Write a Review

Ask friends for feedback

Where Can I Buy It? Compare all Prices

$3.08		amazon.com Marketplace
$3.39	~~5.99~~	LuckyVitamin.com
$4.10	~~5.59~~	Vitacost

Quick, Easy, Effective Relief When Your All Stopped Up!

Mar 14 '07

by batraylover

Pros: Effective; safe to use; affordable; quick relief

Cons: None that I can think of.

Summary: If you have never experienced constipation, I hope that you never will. However, more than likely at some point in your life you will experience constipation. It may be because you ate something that caused your system to "back up", it could be from a ... read more

One Question Site Survey

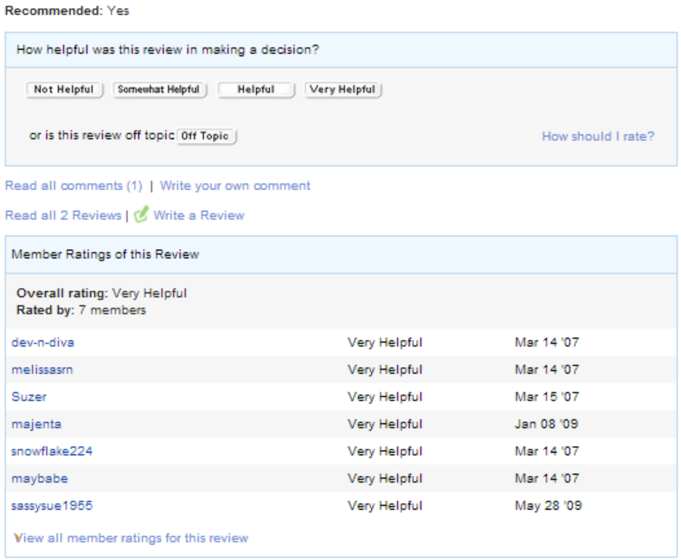
IT TAKES ONLY SECONDS TO ANSWER BELOW

How likely are you to consider purchasing a Panasonic Lumix digital camera?

SELECT ONE ANSWER:

○ Definitely will
○ Probably will
○ May or may not
○ Probably will not
○ Definitely will not

Figure 26 The example of a review page and product we collected



Recommended: Yes

How helpful was this review in making a decision?

or is this review off topic

How should I rate?

Read all comments (1) | Write your own comment

Read all 2 Reviews | Write a Review

Member Ratings of this Review

Overall rating: Very Helpful
Rated by: 7 members

dev-n-diva	Very Helpful	Mar 14 '07
melissasrn	Very Helpful	Mar 14 '07
Suzer	Very Helpful	Mar 15 '07
majenta	Very Helpful	Jan 08 '09
snowflake224	Very Helpful	Mar 14 '07
maybabe	Very Helpful	Mar 14 '07
sassysue1955	Very Helpful	May 28 '09

View all member ratings for this review

Figure 27 The example of a rating page and product we collected