Thesis for the Degree of Doctor of Philosophy

# Improved Trust-Aware Recommender System using Small-Worldness of Trust Networks

Weiwei Yuan

Department of Computer Engineering
Graduate School
Kyung Hee University
Seoul, Korea

August 2010

# Improved Trust-Aware Recommender System using Small-Worldness of Trust Networks

by

Weiwei Yuan

Supervised by

Prof. Young-Koo Lee, Ph.D.

Department of Computer Engineering
Graduate School
Kyung Hee University
Seoul, Korea

August 2010

# Improved Trust-Aware Recommender System using Small-Worldness of Trust Networks

Weiwei Yuan

Submitted to
The Faculty of the Graduate School of Computer Engineering
in Partial Fulfillment of the Requirements
of the Degree of
**Ph.D.**

Thesis Committee:

______________________________
Professor Chae, Ok-Sam

______________________________
Professor Huh, Eui-Nam

______________________________
Professor J. d'Auriol, Brian

______________________________
Professor Lee, Sungyoung

______________________________
Professor Lee, Young-Koo

*Dedicated to my Family*

# Acknowledgement

Many different people provided help, support, and input that brought this thesis to fruition. Firstly, I would like to express my gratitude to my supervisor, Prof. Young-Koo Lee, who gave me the freedom to try out new ideas and gave me continuous support during the research.

I appreciate all those who have helped me in my study, including the professors, my colleges and my friends. I am also in debt of my thesis committee whose comments helped me to very much improve the presentation of the thesis.

I would like to thank my parents for supporting my decision to continue this study. Their love inspires me to be strong whenever I meet the difficulties.

Last and most importantly, I would like to express my highest appreciation to my husband: Dr. Donghai Guan. It is impossible for me to complete this thesis without his help. He shared all my pains, sorrows and depressions in the most difficult time of my research. He continuously encourages me not only in my research but also in my life. I am so grateful to have him as my lifelong partner.

Weiwei Yuan

August 2010

Seoul, Korea

# Abstract

Trust is the measure of willingness to believe in a user based on its competence and behavior within a specific context at a given time. Based on the active users' trusts on the recommenders, the Trust-Aware Recommender System (TARS) suggests the worthwhile information to the users. TARS has superior rating prediction coverage than the traditional recommender system by taking advantages of the trust's transitive property.

The conventional TARS model suffers from several problems. Firstly, it is not optimized. This is because the structure of the dynamic trust model is unknown. The computational complexity of the conventional TARS model can be exponentially more expensive by achieving similar rating predication accuracy and rating prediction coverage. The rating prediction coverage of the conventional TARS model can also be significantly worse by achieving similar rating predication accuracy. Secondly, the conventional TARS model is only effective with the explicit trust statements, which means all the users should explicitly point out their opinions on other users. Since the explicit trusts need extra user effort, they are not always available in the practical recommender systems.

This work first verifies that the trust network used in TARS has the small-world topology. Its small-worldness is independent of its dynamics. Taking advantages of the small-world properties of the trust network, this work optimizes the conventional TARS model to achieve the maximum rating prediction accuracy and the maximum rating prediction coverage with the minimum computational complexity. Secondly, this work improves the conventional TARS model to predict the ratings without the explicit trust statements. This is achieved by generating the implicit trust networks for TARS by the easy available trust sensitive information. Specifically, this work uses the user similarities to get the implicit trusts. By verifying the small-worldness of the implicit trust network, this work applies the small-worldness of the implicit trust network in the rating prediction mechanism of the improved TARS model. The simulation results show that the improved implicit trust based TARS model has high rating prediction accuracy and high rating prediction coverage.

# Table of Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Improved trust-aware recommender system

The Trust-Aware Recommender System (TARS) is the recommender system that suggests the worthwhile information to the users on the basis of trust. Trust is the measure of willingness to believe in a user based on its competence (e.g. goodness, strength, ability) and behavior within a specific context at a given time. It is a directional relationship from the trustor – the user that evaluates its trust on the target user – to the trustee – the user that is the target of the trust evaluation.

TARS has recently been proposed for use since it is able to solve the well-known data sparseness problem of the collaborative filtering (CF) [1, 2]. This is because trust is transitive: it means, if A trusts B and B trusts C, A will trust C to some extend, so even if there is no direct trust between the active users and the recommenders, the active users can build up some indirect trust relationships with the recommenders via the trust propagations. This contributes to the high rating prediction coverage of TARS. Moreover, the rating prediction accuracy of TARS is no worse than the classical CF [1].

The conventional TARS model [1-7] has some shortcomings. Firstly, it is not optimized: its computational complexity can be exponentially more expensive by achieving similar rating predication accuracy and rating prediction coverage, and its rating prediction coverage can be significantly worse by achieving similar rating predication accuracy. Secondly, the conventional TARS models focus on using the explicit trust. That is, the trust should be explicitly pointed out by the users. For examples, the model in [3] requires the users to add those whose ratings they have consistently found to be valuable in their web of trust. The users' trusts on those who are in their web of trust are assigned as 1, and the users' trusts on other users are assigned as 0. These explicit trust statements are then used as the inputs of TARS with the ratings on the items to predict the ratings.

This work motives to optimize the conventional TARS model as well as improve the conventional TARS models to predict ratings without the explicit trust statements. The reason that the conventional TARS model is not optimized is that the structure of the trust network is unknown. The trust network is highly dynamic: any user can join at any time by stating its trust on any existing user of the trust network. This irregular growth leads to the complex structure of the trust network. This work first verifies the small-worldness of the trust network and takes the advantages of the properties of small-world network to optimize the conventional TARS model. Furthermore, since the explicit trust statements are not always available in the practical recommender systems, this work uses other cheap and easy available trust sensitive information to generate the implicit trust for TARS.

The contributions of this work are mainly in three aspects:

--This work verifies that the trust network is the small-world network, and its small-worldness is independent of dynamics. It is shown that the nodes of the trust network are highly clustered, while the distance between two randomly selected nodes is short. This can facilitate the usage of the trust network in various trust-aware applications.

--This work optimizes the conventional TARS model to achieve the maximum rating prediction accuracy and the maximum rating prediction coverage with the minimum computational complexity. This is achieved by optimizing the conventional TARS model to use the small-worldness of the trust networks.

--This work improves the conventional TARS model to predict the ratings without the explicit trust statements. The easy available user similarity information is used to generate the implicit trust for TARS. Based on the small-worldness of the implicit trust network, this work improves the conventional TARS model by predicting ratings only with the ratings. The improved model has high rating prediction accuracy and reasonable rating prediction coverage, which is much higher than the traditional recommender systems.

## 1.2 Thesis outline

Below is a summary of the rest of the thesis:

**Chapter 2:** Related works. This chapter introduces the related works of the trust-aware recommender systems.

**Chapter 3:** Small-world topology of trust networks. This chapter verifies the small-worldness of the trust network, which is independent of its dynamics.

**Chapter 4:** Improved TARS using explicit trust networks. This chapter optimizes the conventional TARS model by using the small-worldness of the trust network.

**Chapter 5:** Improved TARS using implicit trust networks. This chapter generates the implicit trust network by the user similarity, and improves the conventional TARS model by the small-worldness of the implicit trust network.

**Chapter 6:** Conclusions and future works. This chapter summarizes the main contributions of this work and discusses the future research directions for the work presented in this thesis.

# Chapter 2

# Related works

## 2.1 Using trust in recommender system

Collaborative Filtering (CF) [34, 35, 36, 37] collects opinions from users in the form of ratings on items. The recommendations produced are based only on the opinions of users similar to the current user (neighbours). The advantage over content-based RS is that the algorithm doesn't need a representation of the items in terms of features but it is based only on the judgments of the user community.

Collaborative Filtering stresses the concept of community, where every user contributes with her ratings to the overall performances of the system [38, 39, 40]. The traditional input to a CF algorithm is a matrix in which rows represents users and columns items. The entry at each element of the matrix is the user's rating of that item. CF performs three steps:

- It compares the current user's ratings against every other user's ratings. CF computes a similarity value for every other user, where 1 means totally similar and -1 totally dissimilar. Usually the similarity measure is the Pearson correlation coefficient, but any other could be used [7]. The coefficient is computable only if there are items in common rated by both users. If this situation does not occur (as it is often the case), two users are not comparable.
- Based on the ratings of the most similar users (neighbours), it predicts the rating the current user would give to every item she has not yet rated.
- It suggests to the user the items with highest predicted rating.

The standard CF schema is simple but very effective, however it has some weaknesses. The CF algorithm is typical of a lazy, instance based learning algorithm. Such algorithms suffer can be computationally very expensive at query time, since they need search all the user profiles to find the best set of neighbours. This problem means that current RS cannot scale to large environments with millions of users and billions of items (for example, the envisioned Semantic Web [1]). This is also a very slow step, in the sense it can takes from some seconds up to some minutes to find neighbours of one user. For this reason, it is not feasible to do it when a recommendation request is made by the user and hence this should be done periodically offline. However this means that recommendations are not always up to date and that user ratings do not take effect immediately.

User similarity [41, 42, 43] is computable only against few users. The first step suffers another problem. In order to be able to create good quality recommendations, RSs should be able to compare the current user against every other user with the goal of selecting the best neighbours with the more relevant item ratings. This step is mandatory and its accuracy affects the overall system accuracy: failing in finding “good" neighbours will lead to poor quality recommendations. However, since the ratings matrix is usually very sparse because users tend to rate few of the millions of items, it is often the case that two user don't share the minimum number of items rated in commons required by user similarity metrics for computing similarity. For this reason, the system is forced to choose neighbours in the small portion of comparable users and is probably going to miss other non-comparable but relevant users. Usually this does not happen for users with hundreds of ratings but for users with few ratings. However it can be argued that it is more important (and hard) for an RS to provide a good recommendation to a user with few ratings in order to invite her to provide more ratings and keep using the system than to a user with many ratings that is probably already using the system regularly.

Easy attacks by malicious insiders [44, 45, 46]. Recommender Systems are often used in e-commerce sites (for example, in Amazon.com). In those contexts, being able to influence recommendations could be very attractive: imagine if someone could \force" Amazon.com to always recommend the book she wrote. However, subverting standard CF techniques is very easy [10]. The simplest attack is the copy-profile attack: the attacker can copy the ratings of target user and the system will think the attacker is the most similar user to target user. In this way every additional item the attacker rates highly will probably be recommended to the target user. Since currently RSs are mainly centralized servers, creating a “fake" identity is a time-consuming activity and hence these attacks are not currently heavily carried on and studied. However we believe that, as soon as the publishing of ratings and opinions becomes more decentralized (for example, with SemanticWeb formats such as RVW [2] or FOAF [3]), these types of attacks will become more and more an issue. Basically, creating such attacks will become as widespread as spam is today, or at least as easy.

To solve the problems of the conventional CF, a number of researches [1, 6, 7, 8, 9] have focused on extending the recommender system with the trust-awareness.

## 2.2 Conventional TARS model

To solve the problems of the conventional CF, a number of researches [1, 6-9, 55-78] have focused on extending the recommender system with the trust-awareness. Among these works, the TARS model proposed by Massa and Avesani [1, 2, 10, 11] is the most popular one. In addition, their model has already been used in a practical application named Moleskiing.it [12]. Due to its popularity, their TARS model is used as the basis of analysis in this research. The conventional TARS model specifically refers to their model in this research.

The architecture of TARS [1] is shown in Fig. 1. The inputs are the trust matrix and the rating matrix. The output is the predicted ratings on the items for different users. The trust matrix is the collection of the trust relations between the users of the recommender system. Each element of the trust matrix describes the trust between two users. The rating matrix records the users’ ratings on the items. Each element of the rating matrix is the rating given by a user on a particular item.

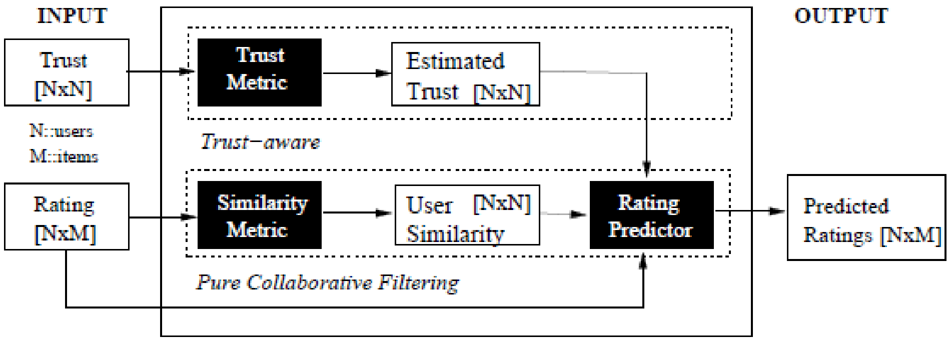

**Fig. 1.** Trust-aware recommender system architecture

**Table 1** Notations used in the TARS

| Symbol | Explanation |
|---|---|
| $i$ | Item |
| $a$ | Active user |
| $u$ | Recommender |
| $\overline{r_a}$ | Average rating of the active user |
| $\overline{r_u}$ | Average rating of the recommender |
| $r_{u,i}$ | Recommender’s rating on the item |
| $w_{a,u}$ | Active user’s weight to the recommender |
| $p_{a,i}$ | Predicted rating for the active user on the item |

The rating prediction mechanism of the conventional TARS model is similar as that of CF. The difference is that CF weights each recommendation based on the active user's similarity with the recommender, while TARS weights each recommendation based on the active user's trust on the recommender:

$$p_{a,i} = \overline{r_a} + \frac{\sum_{u=1}^{k} w_{a,u}(r_{u,i} - \overline{r_u})}{\sum_{u=1}^{k} w_{a,u}} . \tag{1}$$

$w_{a,u}$ is calculated as:

$$w_{a,u} = \frac{d_{\max} - d_{a,u} + 1}{d_{\max}} , \tag{2}$$

where $d_{\max}$ is the maximum allowable propagation distance (MAPD) between users of the recommender system. The value of MAPD is preset by the administrator of TARS. $d_{a,u}$ is the active user *a*'s trust propagation distance to the recommender *u*. In TARS, the trust propagation distance refers to the number of hops in the shortest trust propagation path from the trustor to the trustee.

As shown in the prediction mechanism of the conventional TARS model, MAPD is the fundamental parameter for the rating prediction. However, existing works of TARS did not propose any mechanism to set MAPD. They just randomly choose some value for this extremely important parameter. For example, in [1], the authors randomly set the value of MAPD as 1, 2, 3 and 4 to conduct different experiments of TARS. They did not verify whether these values are the suitable values. And they did not consider the relationship between the value of MAPD and the scale of TARS. On one hand, if the value of MAPD is set too small, TARS might lose some valuable recommendations. On the other hand, the computational complexity of constructing trust networks for TARS is $O(k^{d_{\max}})$ , in which $k$ is the number of trusts stated per user, and $d_{\max}$ is the value of MAPD, so if the value of MAPD is set too big, the computational complexity of TARS increases exponentially. Intuitively, the optimized value of MAPD for TARS should have some relationship with the topology of the trust network. This work therefore analyzes the characteristics of the trust network and optimizes the conventional TARS model based on the topology of the trust network.

# Chapter 3

# Small-world topology of trust networks

The trust network has been widely used in many applications [1], such as the recommender system [2, 3] and the security mechanism [4]. Despite of its popularity, little is known about its topology. This is because the trust network is highly dynamic: a user can join at anytime by stating its trust on any existing user. This irregular growth leads to the complex structure of the trust network. In essence, the topology of the trust network is the important information to optimize its usage in the trust-aware applications, so it is essential to make clear its structure. Since some complex networks, such as the World Wide Web [5] and the e-mail network [6], have been verified to have the small-world topology, some works assume that the trust network also has the small-world nature. These works include, for instance, the trust-based security mechanism [7], the trust-based multiagent system [8] and the trust network modeling [9].

Though the trust network has been assumed to have the small-world topology by the existing works, to the best of my knowledge, no one has proved its small-worldness experimentally or theoretically. By analyzing the trust networks extracted from five public released datasets, this work contribute to verify that the trust network has the small-world topology: on one hand, the nodes of the trust network are highly clustered, which is similar to the regular network; on the other hand, the distance between two randomly selected nodes is short, which is similar to the random network. Further analysis shows that the small-worldness of the trust network is independent of its dynamics.

## 3.1 Introduction of small-world networks

The small-world network is a kind of network between the regular network and the random network. The regular network is highly clustered yet has long distance between two randomly selected nodes. The random network is not clustered yet has short distances between nodes. The small-world network is defined as the network that has [13]: (1) large clustering coefficient, which is much larger than that of its corresponding random network, and (2) short average path length, which is almost as short as that of its corresponding random network, in which a network's corresponding random network refers to the random network that has the same number of nodes and same number of edges per node as this network. The relationship between the regular network, the random network and the small-world network is summarized in Fig. 2. The explanations of the notations used in this chapter are listed in Table 2.

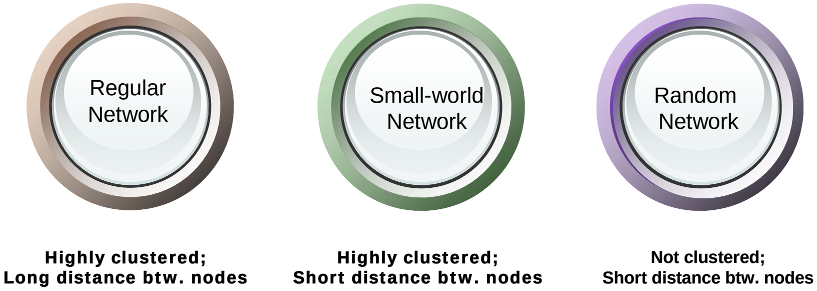
Regular Network
Small-world Network
Random Network
Highly clustered; Long distance btw. nodes
Highly clustered; Short distance btw. nodes
Not clustered; Short distance btw. nodes

**Fig. 2.** Comparison between the regular network, the random network and the small-world network

**Table 2** Notations used in the small-worldness verification

| Symbol | Explanation |
|---|---|
| $n$ | Size of the network |
| $k$ | Average degree of the nodes in the network |
| $k_i$ | Degree of node $i$ |
| $C_i$ | Clustering coefficient of node $i$ |
| $C$ | Clustering coefficient of the network |

| | |
|---|---|
| $C^R$ | Clustering coefficient of the random network |
| $L$ | Average path length of the network |
| $L^R$ | Average path length of the random network |

The clustering coefficient represents the cliquishness of a typical neighborhood [13], i.e., how close the node and its neighbors are to be a complete network. The clustering coefficient of a network is the mean of the clustering coefficient of each node, in which the clustering coefficient of a node is the fraction of the allowable edges and the edges that actually exist between the neighbors of this node [13]:

$$C = \frac{1}{n}\sum_{i=1}^{n} C_i = \frac{1}{n}\sum_{i=1}^{n} \frac{\text{(number of connected neighbor pairs)}}{k_i(k_i - 1)} . \qquad (3)$$

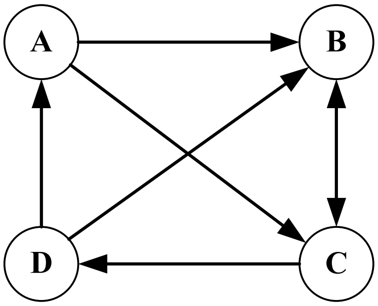

**Fig. 3.** A network with 4 nodes and 7 edges

The network shown in Fig. 3 is used as an example to explain the calculation of equation

(1). Node $A$ has 3 neighbors, i.e. $B$, $C$ and $D$, so at most 6 edges can exist between $A$ 's neighbors. Four edges actually exist in $A$ 's neighborhood: $BC$ , $CB$ , $CD$ and $DB$ . So $C_A$ = 4/6 = 2/3, and similarly $C_B$ = 1/2,

$C_C$ = 1/2 and $C_D$ = 2/3. The clustering coefficient of the network is: $C$ = ($C_A + C_B + C_C + C_D$)/4 = 7/12.

The clustering coefficient of a random network with $n$ nodes and $k$ edges per node is calculated as [13]:

$$C^R = \frac{k}{n}. \tag{4}$$

The average path length $L$ is defined as the number of edges in the shortest path between two nodes, averaged over all pairs of nodes [13]. The average path length of a random network with $n$ nodes and $k$ edges per node is calculated as [13]:

$$L^R = \frac{\ln(n)}{\ln(k)}. \tag{5}$$

## 3.2 Experimental verifications on the small-worldness of trust networks

This work experimentally verifies the small-worldness of the trust networks using data extracted from the real applications. The experimental verification methodology is used since it is the most popular way to verify the small-world topology of various networks [13-18].

### 3.2.1 Experimental setup

The properties of five trust networks are examined to verify the small-worldness. These trust networks are extracted from five public released datasets respectively. These datasets are the Epinions dataset, the Kaitiaki dataset, the Squeakfoundation dataset, the Robots dataset and the Advogato dataset. These datasets are chosen since they are all the public available datasets when this research began. They are available at trustlet.org[1].

Epinions consists of 49288 users and 487183 trust statements. The data is extracted from epinions.com [2] from November to December of 2003. Epinions.com is a recommender system that recommends items based on other users' ratings. In addition to the ratings on the items, the users are required to explicitly express their trust on other users. The trustor evaluates its trust on the trustee as 1 if the trustor consistently finds the ratings given by the trustee are valuable, otherwise, the trustor evaluates its trust on the trustee as 0.

Advogato consists of 5412 users and 54012 trust statements. The data is extracted from advogato.org [3] on June 1, 2009. Advogato.org is an online community site dedicated to free software development. On advogato.com users can certify each other as several levels: Observer, Apprentice, Journeyer or Master [19]. Masters are supposed to be excellent programmers who work full time on free software, Journeyers contribute significantly, but not necessarily full-time, Apprentices contribute in some way, but are still acquiring the skills needed to make more significant contributions, and observers are users without trust certification. These certifications are regarded as the trust statements of Advogato.

Kaitiaki consists of 64 users and 154 trust statements. The data is extracted from kaitiaki.org[4] on September 1, 2008. The trust statements of Kaitiaki are weighted at four different levels: Kaitiro, Te Hunga Manuhiri, Te Hunga Käinga, Te Komiti Whakahaere. Squeakfoundation consists of 461 users and

[1] http://www.trustlet.org/wiki/Datasets

[2] http://www.epinions.com/

[3] http://www.advogato.org/

[4] http://www.kaitiaki.co.nz/

2697 trust statements. The data is extracted from squeak.org[5] on November 1, 2008. The trust statements of Squeakfoundation are weighted at three different levels: Apprentice, Journeyer, and Master. Robots consists of 1646 users and 3456 trust statements. The data is extracted from robots.net[6] on March 1, 2009. The trust statements of Robots are weighted at three different levels: Apprentice, Journeyer, and Master. Kaitiaki.org, squeak.org and robots.net are all web community sites which use the same software which powers the Advogato web community site, mod virgule. These three datasets are much smaller than the Advogato dataset.

The characteristics of the explored trust networks are summarized in Table 3. All users involved in these trust networks act as the trustors, the trustees or both.

**Table 3** Description of the trust networks used in this research

| | **Number of nodes** | **Number of edges per node** |
|---|---|---|
| **Epinions** | 49288 | 9.88 |
| **Kaitiaki** | 64 | 2.41 |
| **Squeakfoundation** | 461 | 5.85 |
| **Robots** | 1646 | 2.1 |
| **Advogato** | 5412 | 9.98 |

[5] http://www.squeak.org/Foundation/

[6] http://robots.net/

### 3.2.2 Experimental results

Experiments are held on the above trust networks to verify their small-worldness.

Firstly, this work verifies that the trust networks have large clustering coefficients. Using equation (3) and equation (4), the clustering coefficients of the explored five trust networks and their corresponding random networks are evaluated, which are summarized in Table 4. The detailed distributions of the explored trust networks' clustering coefficients are given in Fig. 4. It shows that: though the clustering coefficients of some users are small (near 0), those of the majority users are greater than 0.1. A portion of the clustering coefficients even equals to 1. This means the neighbors of some users are fully connected. This is very different from the random network. The comparison between the clustering coefficients of the trust networks and those of their corresponding random networks clearly shows that: the trust network has much larger (higher order of magnitude) clustering coefficients than its corresponding random network. This satisfies the first condition of the small-world network's definition.

**Table 4** Clustering coefficients of the trust networks and their corresponding random networks

| | $n$ | $k$ | $C$ | $C^R$ |
|---|---|---|---|---|
| **Epinions** | 49288 | 9.88 | 0.217 | $2\times10^{-4}$ |
| **Kaitiaki** | 64 | 2.41 | 0.24 | $3.77\times10^{-2}$ |
| **Squeakfoundation** | 461 | 5.85 | 0.44 | $1.27\times10^{-2}$ |
| **Robots** | 1646 | 2.1 | 0.22 | $1.28\times10^{-3}$ |
| **Advogato** | 5412 | 9.98 | 0.23 | $1.84\times10^{-3}$ |

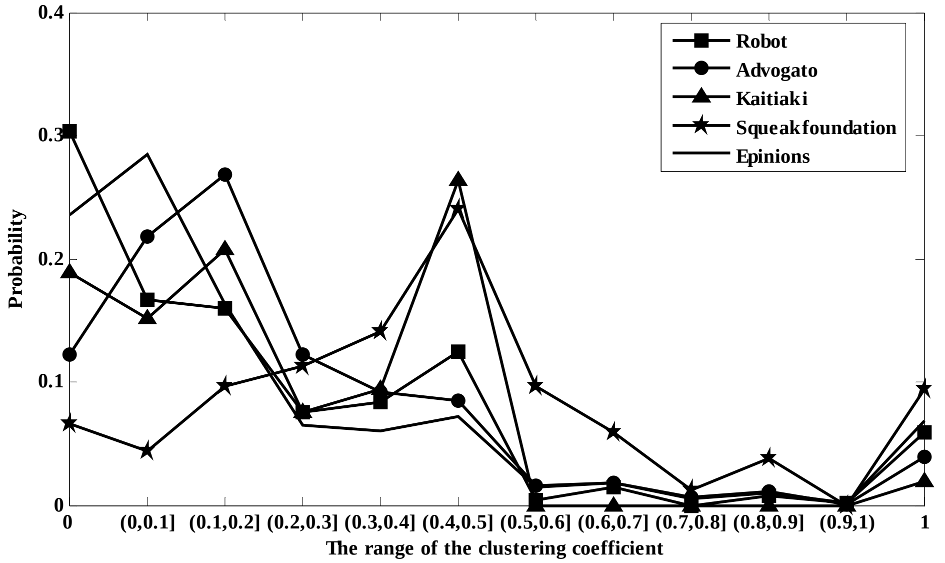

**Fig. 4.** Distribution of the trust networks' clustering coefficients

Secondly, this work verifies that trust networks have short average path lengths. For large networks, measuring all-pair distances is computational expensive, so an accepted procedure is to measure it over a random sample of nodes [20]. The average path lengths for the larger networks (Epinions and Advogato) in Table 3 are measured on a random sample of 5%. The average path lengths for the smaller networks (Kaitiaki, Squeakfoundation and Robots) in Table 3 are measured on all pairs of nodes. The distributions of the five trust networks' average path lengths are given in Fig. 5. It shows that the trust networks have very small number of direct trusts, i.e., where the path length equals to 1. By propagating trust, users can build up their trust relationships with others within several hops. Another important observation is that very small number of the trust propagations has long distance, e.g. the probabilities that the path lengths are longer than 8 hops (if any) are less than 1%. The path lengths of most trust propagations are from 2 hops to 6 hops. In more details: (1) the maximum path length of Epinions is 11 hops, and its average path length is 3.96 hops; (2) the maximum path length of Kaitiaki is 5 hops, and its average path length is 2.16 hops; (3) the maximum path length of Squeakfoundation is 6 hops, and its average path length is 2.85 hops; (4) the maximum path length of Robots is 11 hops, and its average path length is 3.94 hops; (5) the maximum path length of Advogato is 9 hops, and its average path length is 3.8 hops.

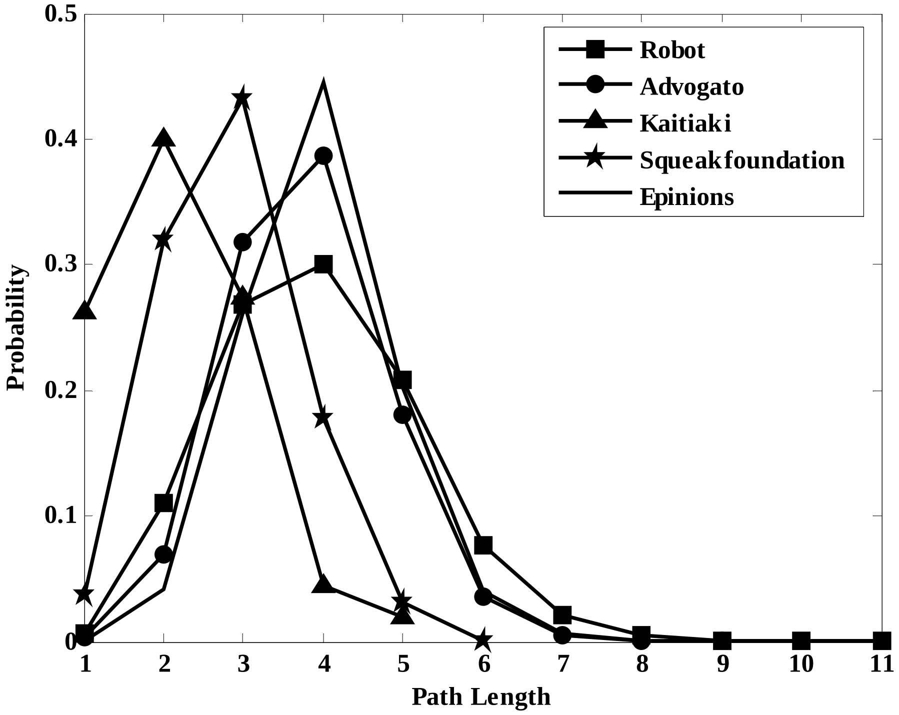

**Fig. 5.** Distribution of the trust networks' path lengths

Using equation (5), the average path lengths of the explored five trust networks' corresponding random networks are evaluated, which are summarized in Table 5. Comparing the average path lengths of the trust networks with those of their corresponding random networks, it is obvious that the trust networks have similar (the same order of magnitude) average path lengths as their corresponding random networks. This satisfies the second condition of the small-world network's definition.

**Table 5** Average path lengths of the trust networks and their corresponding random networks

| | $n$ | $k$ | $L$ | $L^R$ |
|---|---|---|---|---|
| **Epinions** | 49288 | 9.88 | 3.96 | 4.71 |
| **Kaitiaki** | 64 | 2.41 | 2.16 | 4.73 |
| **Squeakfoundation** | 461 | 5.85 | 2.85 | 3.47 |
| **Robots** | 1646 | 2.1 | 3.94 | 9.98 |
| **Advogato** | 5412 | 9.98 | 3.80 | 3.74 |

### 3.2.3 Analysis on the small-worldness of trust networks

Using the above characteristics on the clustering coefficient and the average path length, this work compares the trust networks with some well-known small-world networks documented in the literatures: the World Wide Web [5], the human language network [13], the e-mail network [6], the human brain network [14, 15], the film actors network [11], the power grid network [11], and the C. elegans network [11]. The characteristics of these networks and those of their corresponding random networks are shown in Table 6. Based on Table 6, a further comparison between the small-world characteristics of the trust networks and these networks is presented in Fig. 6. The axes of Fig. 6 represent the ratios of the selected networks and their corresponding random networks. Note that most small-world networks are concentrated around where the average path length ratio equals to 1. This means that the selected networks have similar average path length as their corresponding random networks. In addition, the clustering coefficient ratios of most networks are greater than 10. This means that the selected networks have much larger clustering coefficients than their corresponding random networks. The comparisons of Table 6 and Fig. 6 clearly show that the trust networks have the same properties as other well-known small-world networks: they are highly clustered yet have small average path lengths. This work therefore draws the conclusion that the trust networks are the small-world networks.

**Table 6** Small-world characteristics of some well-known small-world networks

| | $n$ | $k$ | $C$ | $C^R$ | $L$ | $L^R$ |
|---|---|---|---|---|---|---|
| **World Wide Web** | 153127 | 19 | 0.156 | $1.2\times10^{-3}$ | 4.06 | 4.048 |
| **Human language** | 460902 | 70.79 | 0.437 | $1.55\times10^{-4}$ | 2.67 | 3.06 |
| **E-mail network** | 56969 | 2.95 | 0.03 | $4.82\times10^{-5}$ | 4.95 | 10.1 |
| **Human brain** | 90 | 4.5 | 0.53 | 0.05 | 2.49 | 2.99 |
| **Film actors** | 225226 | 61 | 0.79 | $2.7\times10^{-4}$ | 3.65 | 2.99 |
| **Power grid** | 4941 | 2.67 | 0.8 | $5\times10^{-3}$ | 18.7 | 12.4 |
| **C. elegans** | 282 | 14 | 0.28 | 0.05 | 2.65 | 2.25 |

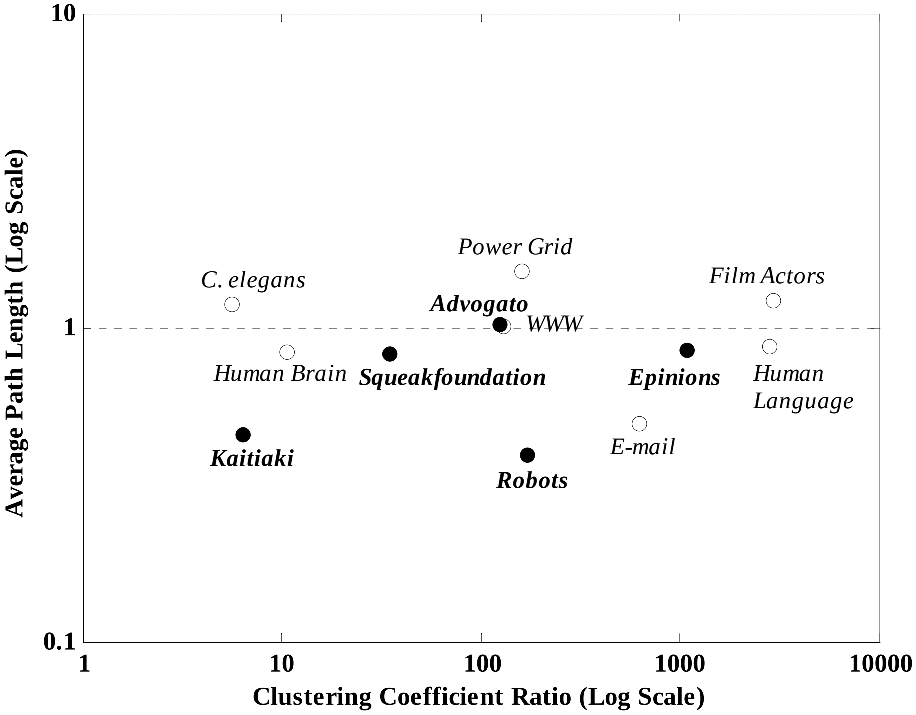

**Fig. 6.** Small-world characteristics of the trust networks and some well-known small-world networks

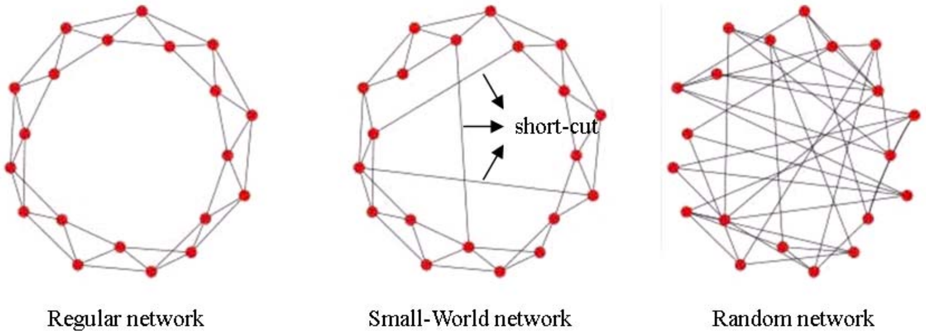

**Fig. 7.** Explanations of the small-worldness of the trust networ

The small-worldness of the trust network results from the existence of some long range edges which connects different subgroups of the trust network, as shown in Fig. 7. These long range edges act as the short-cut between users.

Because of these “short-cuts”, the users in one group can easily reach the users of another groups, this contributes to the short trust distances between users. For the regular network, since there is not such long-range edge, its average path length is long. For the random network, since there exists a number of long-range edges, its average path length is also short, similar as the small-world network.

## 3.3 The small-worldness of dynamic trust networks

The experiments in Chapter 3.2 verify the small-worldness of the trust network. The method this work uses is the one used in the small-world verification of all other networks: to show the small-worldness of a network, the conventional method verifies that the network has large clustering coefficient and small average path length. The verifications are held on the data extracted from the objective network. The experimental data used for the verifications are the static data, i.e., they only reflect of the status of the network at one moment. So the conventional method only verifies the small-worldness of the static network. However, since some networks, such as the trust networks, are dynamically changing, further verifications should be held on the small-worldness of the networks in dynamics. This work achieves this by verifying the small-worldness of the dynamic trust network via verifying its scale-freeness. The scale-free network is a kind of network whose degree distribution decays as a power law [17]. It is one kind of the small-world network [16, 17]. Many large-scale complex networks are scale-free [18]. The relationship between the small-world network and the scale-free network is given in Fig. 8, in which the broad-scale network is characterized by a degree distribution that has a power law regime followed by a sharp cutoff and the single-scale network is characterized by a degree distribution with a fast decaying tail [16].

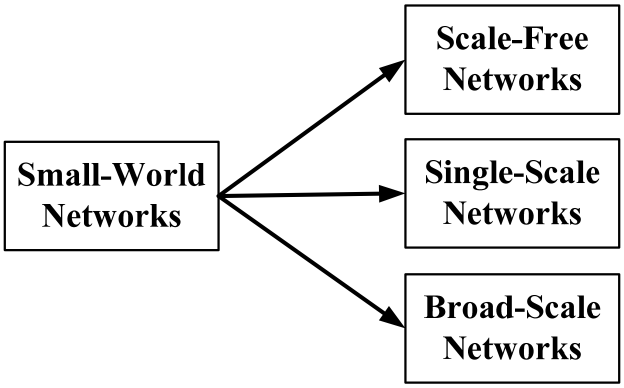

**Fig. 8.** Relationship between the small-world network and the scale-free network

The scale-freeness of a network ensures that this network still has the scale-free structure in dynamics. This is because the scale-free structure of such network is independent of its scale [19]. There are some highly connected nodes in the scale-free network, dominating the connectivity. Unlike the random networks, the probability with which a new node connects to the existing nodes is not uniform in the scale-free network. There is a higher probability that it will be linked to a node that already has a large number of connections [19]. This contributes to the network’s continuous scale-freeness when the network changes.

Since the scale-free network is a kind of the small-world network, if we can verify that the trust network is the scale-free network by the static network data, we can draw the conclusion that the trust network is the small-world network. Moreover, since the scale-freeness of the network is independent of its dynamics, we can further make the conclusion that the dynamically changing trust network is the small-world network. This verification method only uses the static trust network data. Extra data that describe the status of the trust networks in dynamics are not needed.

In addition to its ability in verifying the small-worldness of the dynamic trust networks, verifying the scale-freeness is computational less expensive. The conventional method needs to calculate the clustering coefficient and the average path length of the trust network respectively. The clustering coefficient of a network is the mean of the clustering coefficient of each node, in which the clustering coefficient of a node is the fraction of the allowable edges and the edges that actually exist between the neighbors of this node [10]. To calculate the clustering coefficient, the conventional method needs to make clear the connections between all pairs of nodes in each node's neighborhood. The average path length is the number of edges in the shortest path between two nodes, averaged over all pairs of nodes [10]. To calculate the average path length, the conventional method needs to make clear the trust propagation distance between any two nodes of the trust networks. However, to verify the scale-freeness of the trust network, we only need to calculate the degree distributions of each node. That is, we only need to know the direct trust between the nodes of the trust network, while we do not need to know the trust propagation relationships between these nodes.

The degree distributions of the trust networks shown in Table 3 are examined to verify their small-worldness via the scale-freeness. The trust is asymmetrical, i.e., if A trusts B, B does not necessarily need to trust A. So the trust network is the directed network. This work therefore distinguishes the indegree distribution and the outdegree distribution of the trust networks. The degree distributions of the explored five trust networks are presented in Fig. 9 ~ Fig. 18. Note that some parts of axes in the figures are marked as 0(0.1). This is because the indegree or outdegree of some nodes equals to 0, but 0 is not a valid value for the logarithm. To show the degree distributions of these nodes, this work uses 0.1 to approximately substitute 0 when calculating the logarithm of the degrees.

It is clearly shown in the experimental results that the nodes' indegree distribution and outdegree distribution both follow the power-law in each trust network. That is, the degree distributions follow the rule $P(k) \sim k^{-\gamma}$ , in which $P(k)$ is the probability that a randomly selected node has exactly $k$ edges, and $\gamma$ is the power of the degree distributions. The powers of the explored trust networks' degree distributions are further listed in Table 7, in which $\gamma_{in}$ and $\gamma_{out}$ represent the power of the indegree distribution and the power of the outdegree distribution respectively in Fig. 9 ~ Fig. 18. This work therefore makes the conclusion that the trust networks are the scale-free networks according to the definition of the scale-free networks. This work makes the further conclusion that the dynamic trust networks are the small-world networks.

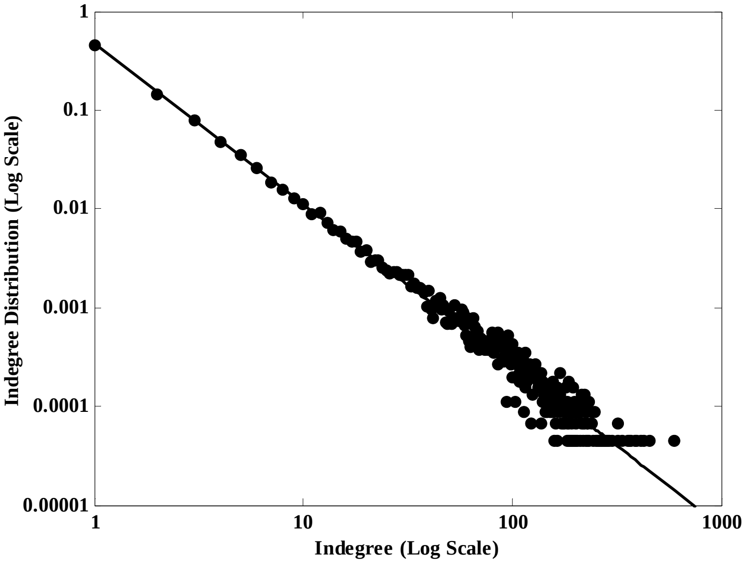

**Fig. 9.** Indegree distribution of Epinions

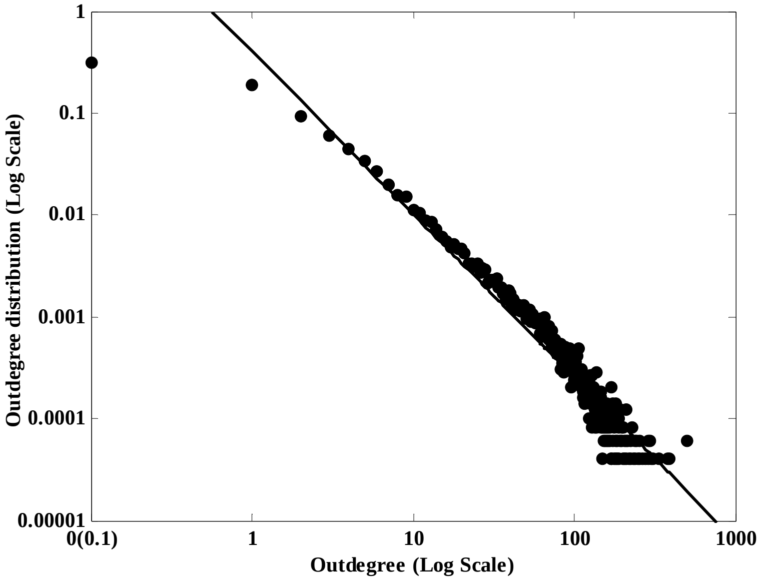

**Fig. 10.** Outdegree distribution of Epinions

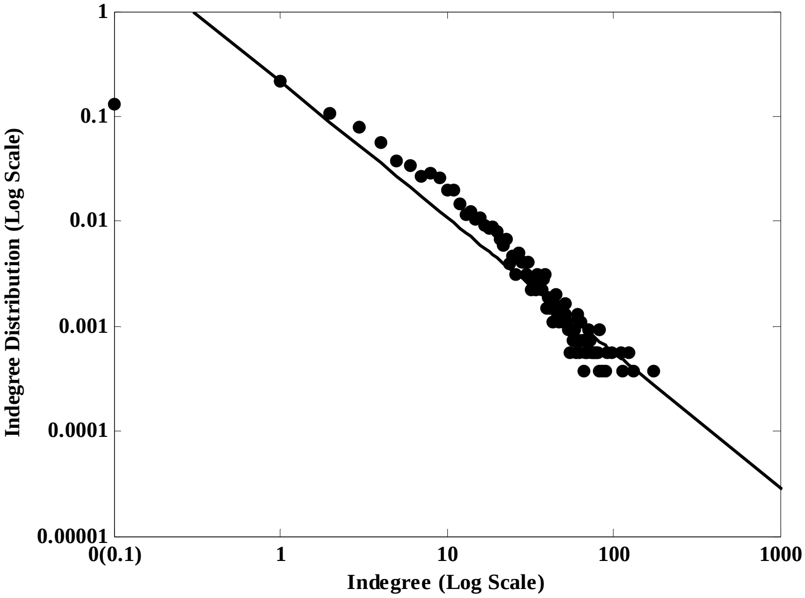

**Fig. 11.** Indegree distribution of Advogato

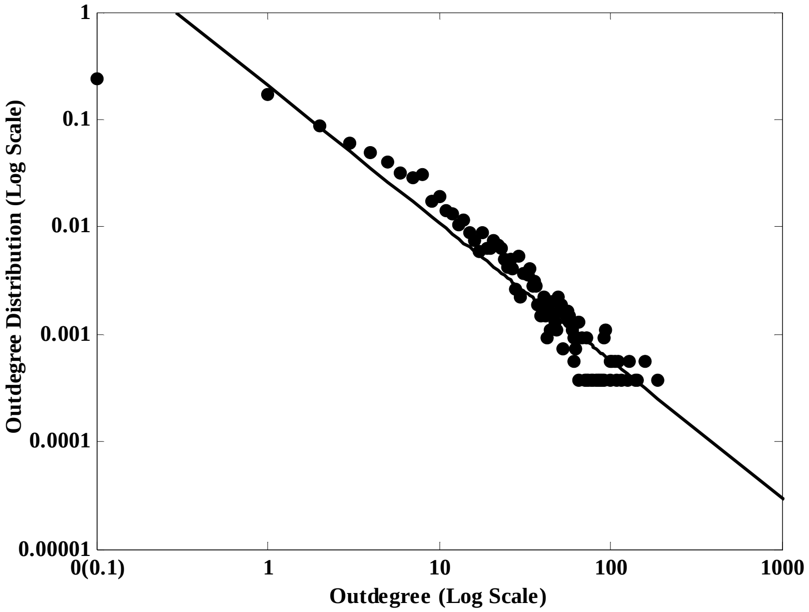

**Fig. 12.** Outdegree distribution of Advogato

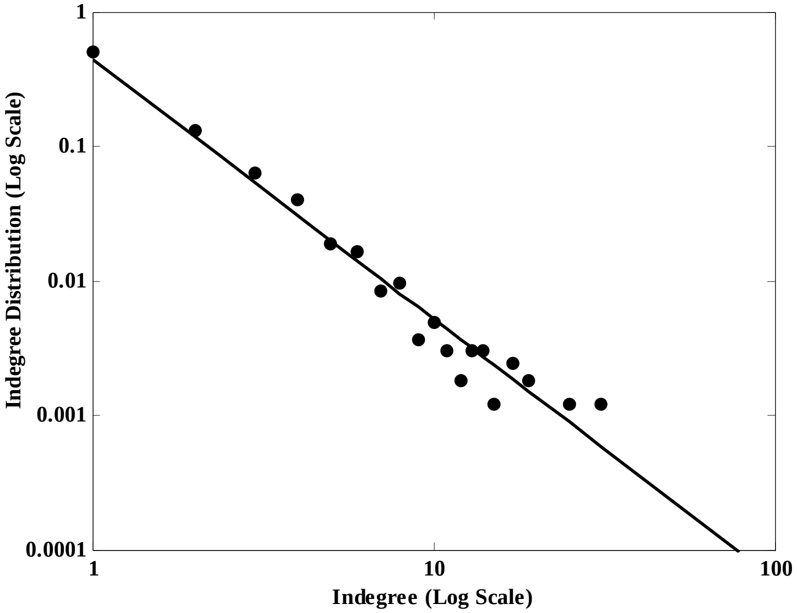

**Fig. 13.** Indegree distribution of Robots

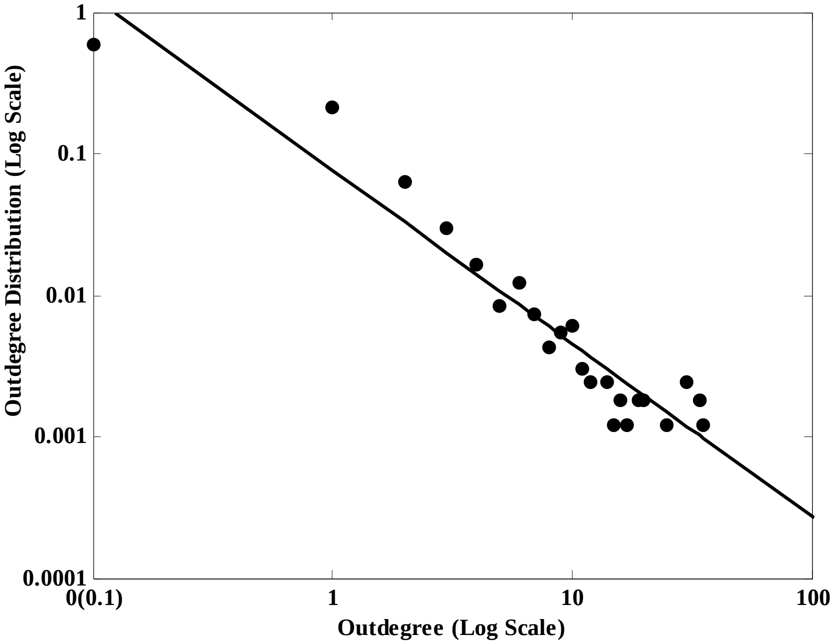

**Fig. 14.** Outdegree distribution of Robots

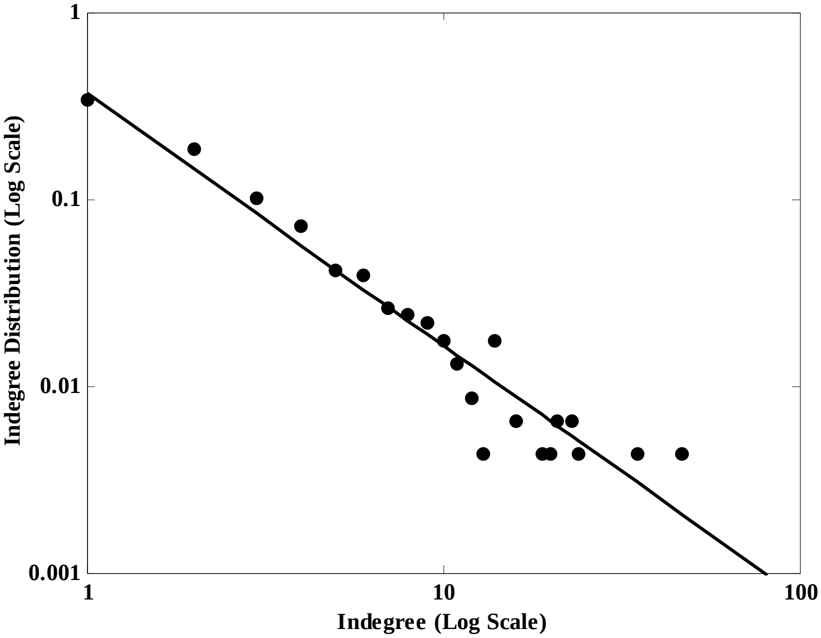

**Fig. 15.** Indegree distribution of Squeakfoundation

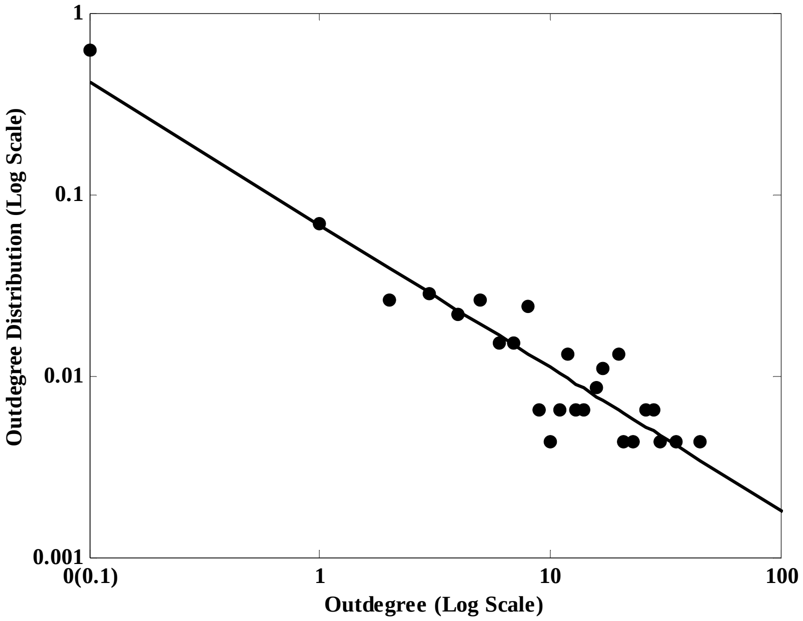

**Fig. 16.** Outdegree distribution of Squeakfoundation

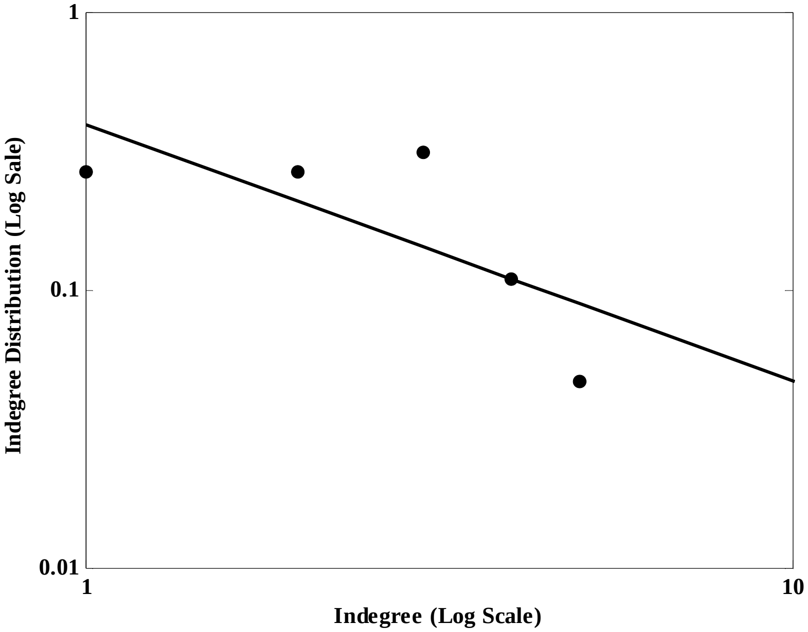

**Fig. 17.** Indegree distribution of Kaitiaki

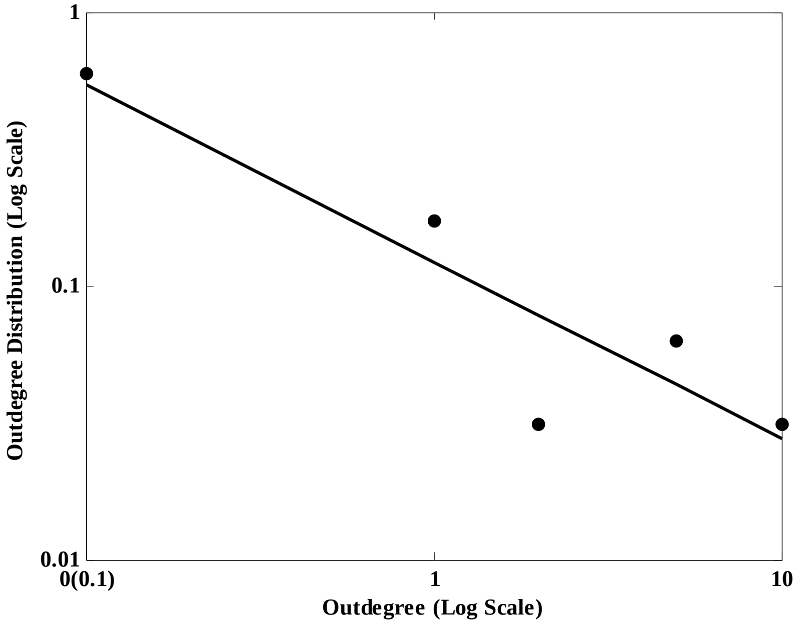

**Fig. 18.** Outdegree distribution of Kaitiaki

**Table 7** Indegree distribution and outdegree distribution of the trust networks

| | *n* | *k* | $\gamma_{in}$ | $\gamma_{out}$ |
|---|---|---|---|---|
| **Epinions** | 49288 | 9.88 | 1.53 | 1.6 |
| **Kaitiaki** | 64 | 2.41 | 0.92 | 0.64 |
| **Squeakfoundation** | 461 | 5.85 | 1.93 | 0.79 |
| **Robots** | 1646 | 2.1 | 1.93 | 1.23 |
| **Advogato** | 5412 | 9.98 | 1.29 | 1.28 |

# Chapter 4

# Improved TARS using explicit trust networks

Despite of its high rating prediction accuracy and high rating prediction coverage, the conventional TARS model suffers from the problem that it is not optimized: its computational complexity can be exponentially more expensive by achieving similar rating predication accuracy and rating prediction coverage, and its rating prediction coverage can be significantly worse by achieving similar rating predication accuracy. This chapter proposes a novel TARS model which can effectively overcome the weakness of the conventional TARS model. This is achieved by leveraging the verified small-worldness of trust networks. Experimental results clearly show that: the proposed model is superior to the conventional one since it is able to achieve the maximum rating prediction accuracy and the maximum rating prediction coverage with the minimum computational complexity

## 4.1 The proposed TARS model

For different sized TARS, it is hard to directly point out the value of MAPD between two randomly selected users. However, since the trust network of TARS is the small-world network, it is easy to get the approximate average trust propagation distance between two randomly selected users of the trust network: it is similar to the average path length of the trust network's corresponding random network. We only need to know the size and the average degrees of the trust network. Since the value of MAPD is unknown and the average path length of the trust network is the only available information about the distance between two users, the proposed rating prediction algorithm heuristically chooses the average path length of the trust network as the value of MAPD for

TARS. The details of the rating prediction algorithm of the proposed TARS model are shown in Table 8.

**Table 8** The proposed rating prediction algorithm

**Algorithm:** The proposed rating prediction algorithm

**Input:** $\mathbf{T}$ (trust matrix), $\mathbf{R}$ (rating matrix)

**Parameter:** $a$ (active user), $i$ (item), $d_{\max}$ (the maximum allowable propagation distance), $n$ (size of the trust network), $k$ (average degrees of the trust network).

**Output:** $p_{a,i}$ ( $a$ 's predicted rating on $i$ )

**Phase 1:** MAPD calculation.

**Phase 2:** Recommender searching.

**Phase 3:** Recommender weighting.

**Phase 4:** Rating calculation.

The proposed TARS model consists of four phases:

The first phase is the MAPD calculation. In this phase, the average path length of the trust network used in TARS is used as the value of MAPD. Due to small-worldness of the trust network, this value approximately equals to the average path length of this trust network's corresponding random network:

$$d_{\max} = \lceil L \rceil \approx \lceil L^R \rceil = \left\lceil \frac{\ln(n)}{\ln(k)} \right\rceil , \tag{6}$$

where $\lceil \cdot \rceil$ represents the ceiling of selected value, e.g. $\lceil L \rceil$ is the ceiling of the average path length of the trust network. The value of $L^R$ is calculated by equation (5). For the Epinions data shown in Table 3, we can get $d_{\max} = \lceil L \rceil \approx \lceil L^R \rceil = \lceil 4.71 \rceil = 5$ for TARS.

The second phase is the recommender searching. In this phase, TARS searches all valid recommenders based on the selected MAPD. A recommender is valid if (1) there is at least one trust propagation path from the active user to the recommender in the trust network, and (2) the trust propagation distance from the active user to the recommender is no longer than $\lceil L \rceil$.

The third phase is the recommender weighting. In this phase, the valid recommenders are weighted based on the relationship between the active users' trust propagation distances to the recommenders and the selected MAPD. This work uses the similar weighting mechanism as the conventional TARS model, as shown in equation (2). The difference is that the proposed model explicitly points out the value of MAPD, which is calculated by equation (7). The weighting mechanism of the proposed model is:

$$w_{a,u} = \frac{\lceil L \rceil - d_{a,u} + 1}{\lceil L \rceil} \approx \frac{\lceil L^R \rceil - d_{a,u} + 1}{\lceil L^R \rceil}, \qquad (7)$$

The last phase is the rating calculation. This phase predicts the ratings by aggregating the recommendations given by the valid recommenders. Each recommendation is weighted with respect to the weight of the recommender, which is calculated by equation (7). The aggregation mechanism used in the proposed model is the same as the conventional TARS model, which is also the one used in CF, as shown in equation (1).

## 4.2 Experimental setup

To verify the performance of the conventional TARS model and the proposed TARS model, this work examines of these models on the data of the Epinions dataset. Data from other datasets used in chapter 3.2.1 are not used to simulate TARS. This is because these datasets only have the trust data while the inputs of TARS need the trust data and the rating data simultaneously.

**Table 9** TARS experimental data

| | | Num of users | Num of items | Num of trusts | Num of ratings |
|---|---|---|---|---|---|
| **Epinions_1** | **Trust Data** | 45275 | - | 461064 | - |
| | **Rating Data** | 31019 | 551392 | - | 8632163 |
| **Epinions_2** | **Trust Data** | 4389 | - | 37843 | - |
| | **Rating Data** | 2275 | 36144 | - | 740422 |
| **Epinions_3** | **Trust Data** | 49288 | - | 487183 | - |
| | **Rating Data** | 20157 | 139633 | - | 664061 |

To provide more evidence on the effectiveness of the proposed method with a single dataset, this work extracted three sets of data from the Epinions dataset based on the timestamp of the trust statements and the ratings. These three sets of data are named as Epinions_1, Epinions_2 and Epinions_3 respectively. Each set of data consists of both the trust data and the rating data. Epinions_1 records trust statements and ratings stated by users in January 2001. Its trust data consists of 45275 users and 461064 trust statements. Its rating data consists of 31019 users' 8632163 ratings on 551392 items. Epinions_2 records trust statements and ratings stated by the users in the year 2002, from January to December. Its trust data consists of 4389 users and 37843 trust statements. Its rating data consists of 2275 users' 740422 ratings on 36144 items. Epinions_3 records trust statements and ratings stated by the users in November and December of 2003. Its trust data is the same as Epinions used in chapter 3.2.1. It consists of 49288 users and 487183 trust statements. The rating data of Epinions_3 consists of 20157 users' 664061 ratings on 139633 items. Both Epinions_1 and Epinions_2 are extracted from the "extended epinions dataset"[7]. Epinions_3 is extracted from the "epinions dataset"[8]. Table 9 is used to summarize these three sets of experimental data. Note that not all users in the trust data are involved in the rating data. This is because some users of the trust network may not give any ratings on the items. E.g. only around 40% users in the trust data of Epinions_3 are involved in the rating data.

## 4.3 Experimental results

This work examines TARS on three aspects to show the effectiveness of the proposed model. These three aspects are the rating prediction accuracy, the rating prediction coverage and the computational complexity.

Using Epinions_1, Epinions_2 and Epinions_3, this work predicts ratings on the rated items of each rating data. Since the scale of each rating data is huge, it is very effort-consuming to predict ratings on all the rated items. This work therefore randomly selects 5% of the rating records from each rating data as the object of the prediction. That is, this work predicts around 400,000 ratings for Epinions_1, around 30,000 ratings for Epinions_2, and around 30,000 ratings for Epinions_3. The MAPD of the proposed rating prediction algorithm is calculated based on the properties of each trust network: for Epinions_1, $d_{\max}$ = $\left\lceil \frac{\ln(45275)}{\ln(461064/45275)} \right\rceil = \lceil 4.62 \rceil$ =5; for Epinions_2, $d_{\max} = \left\lceil \frac{\ln(4389)}{\ln(37843/4389)} \right\rceil$ = $\lceil 3.9 \rceil$ =4, for Epinions_3, $d_{\max} = \left\lceil \frac{\ln(49288)}{\ln(487183/49288)} \right\rceil = \lceil 4.72 \rceil$ =5.

[7] http://www.trustlet.org/wiki/Extended_Epinions_dataset.

[8] http://www.trustlet.org/wiki/Downloaded_Epinions_dataset.

**Table 10** MAE of TARS with different values of MAPD

| | **Epinions_1** | **Epinions_2** | **Epinions _3** |
|---|---|---|---|
| $d_{\max}$ **= 1** | 0.2613 | 0.2155 | 0.8136 |
| $d_{\max}$ **= 2** | 0.2568 | 0.2155 | 0.7542 |
| $d_{\max}$ **= 3** | 0.2576 | 0.2142 | 0.7319 |
| $d_{\max}$ **= 4** | 0.2563 | **0.2139** | 0.7262 |
| $d_{\max}$ **= 5** | **0.2544** | 0.2138 | **0.7253** |
| $d_{\max}$ **= 6** | 0.2546 | 0.2138 | 0.7251 |
| $d_{\max}$ **= 7** | 0.2548 | 0.2138 | 0.7252 |
| $d_{\max}$ **= 8** | 0.2549 | 0.2138 | 0.7253 |
| $d_{\max}$ **= 9** | 0.2550 | 0.2138 | 0.7254 |

The rating prediction accuracy of TARS is measured by the error of the predicted ratings. Specifically, this work calculates the Mean Absolute Error (MAE), since it is very appropriate and useful for evaluating prediction accuracy in offline tests [3]. To calculate MAE, the predicted rating is compared with the real rating and the difference (in absolute value) is the prediction error, this error is then averaged over all predictions to obtain the overall MAE. By predicting the rating on each rated item of Epinions_1, Epinions_2 and Epinions_3, the MAE of TARS with respect to different values of MAPD is reported in Table 10, in which the bold ones are the MAEs calculated by using the proposed method. The experimental results show that: (1) If the value of MAPD is set to be smaller than the suggested value, the rating prediction accuracy of TARS is getting worse. (2) If the value of MAPD is set to be greater than the suggested value, the rating prediction accuracy of TARS dose not change significantly.

The coverage of TARS is measured by both the rating coverage and the recommender coverage. The rating coverage is the portion of items that TARS is able to predict, i.e., the portion of items that the active user can get at least one recommendation. However, this quantity is not always informative about the quality of TARS. TARS is sometimes good on the rating coverage, but only involve small portion of recommenders. This is because an item usually has a number of recommendations, so a good rating coverage does not necessarily imply a good coverage on the recommenders. Since to involve as many recommendations as possible in TARS facilities the rating prediction, this work introduces the term recommender coverage. The recommender coverage is the portion of recommenders that could be involved in TARS. The rating coverage and the recommender coverage of TARS by using different values of MAPD are reported in Table 11 and Table 12 respectively, in which the bold values are the coverage calculated by using the proposed method. The experimental results show that: (1) If the value of MAPD is set to be smaller than the suggested value, both the rating coverage and the recommender coverage of TARS decrease, in which the recommender coverage decreases significantly. (2) If the value of MAPD is set to be greater than the suggested value, the rating coverage and the recommender coverage of TARS do not change significantly. This is because the rating coverage and the recommender coverage are both very high, more than 99%, by using the suggested value of MAPD.

**Table 11** Recommender coverage of TARS with different MAPD

| | **Epinions_1** | **Epinions_2** | **Epinions _3** |
|---|---|---|---|
| $d_{\max}$ **= 1** | 12.52% | 17.92% | 4.10% |
| $d_{\max}$ **= 2** | 74.67% | 87.70% | 30.80% |
| $d_{\max}$ **= 3** | 97.84% | 98.68% | 75.31% |
| $d_{\max}$ **= 4** | 99.80% | **99.85%** | 95.81% |

| | | | |
|---|---|---|---|
| $d_{\max}$ **= 5** | **99.97%** | 99.98% | **99.45%** |
| $d_{\max}$ **= 6** | 100.00% | 100.00% | 99.91% |
| $d_{\max}$ **= 7** | 100.00% | 100.00% | 99.98% |
| $d_{\max}$ **= 8** | 100.00% | 100.00% | 100.00% |
| $d_{\max}$ **= 9** | 100.00% | 100.00% | 100.00% |

**Table 12** Rating coverage of TARS with different MAPD

| | **Epinions_1** | **Epinions_2** | **Epinions _3** |
|---|---|---|---|
| $d_{\max}$ **= 1** | 85.41% | 91.94% | 63.45% |
| $d_{\max}$ **= 2** | 99.29% | 99.70% | 96.52% |
| $d_{\max}$ **= 3** | 99.94% | 100.00% | 99.83% |
| $d_{\max}$ **= 4** | 99.98% | **100.00%** | 100.00% |
| $d_{\max}$ **= 5** | **100.00%** | 100.00% | **100.00%** |
| $d_{\max}$ **= 6** | 100.00% | 100.00% | 100.00% |
| $d_{\max}$ **= 7** | 100.00% | 100.00% | 100.00% |
| $d_{\max}$ **= 8** | 100.00% | 100.00% | 100.00% |
| $d_{\max}$ **= 9** | 100.00% | 100.00% | 100.00% |

The computational complexity of constructing the trust network for TARS is $O(k^{d_{\max}})$, in which $k$ is the number of edges per node in the trust network, and $d_{\max}$ is the value of MAPD. Therefore, if the value of MAPD is set to be smaller than the suggested value, the computational complexity of constructing trust networks for TARS is exponentially less expensive. On the other hand, if the value of MAPD is set to be greater than the suggested value, the computational complexity of constructing trust networks for TARS is exponentially more expensive.

To sum up, though setting the value of MAPD smaller than the suggested value is computational less expensive, the accuracy and the coverage of TARS are worse; while setting the value of MAPD greater than the suggested value leads to similar accuracy and similar coverage of TARS, but it is computational exponentially more expensive. This work therefore draws the conclusion that $\lceil L \rceil$ is a suitable value of MAPD for TARS. This verifies the effectiveness of the proposed method.

Note that $\lceil L \rceil$ is only similar to the average trust propagation distance between two randomly selected users of the trust network, but the experiments show that $\lceil L \rceil$ is a appropriate value of MAPD for TARS. This is because it is the average trust propagation distance between all pairs of users that $\lceil L \rceil$ is similar to. However, not all users are recommenders. Further analysis on the distribution of the average path length between the trustors and the recommenders, which is shown in Fig. 19, shows that: compared with the distribution of the average path length between all pairs of users in the trust network, as shown in Fig. 20, the average path length between the trustors and recommenders are much smaller than that between all pairs of users, and the maximum distance between the trustors and the recommenders are always shorter than that between all pairs of users. This indicates that compared with the non-recommenders or the non-active recommenders, the recommenders tend to have shorter distances with the trustors. This contributes to the effectiveness of the method by setting $\lceil L \rceil$ as the value of MAPD for TARS.

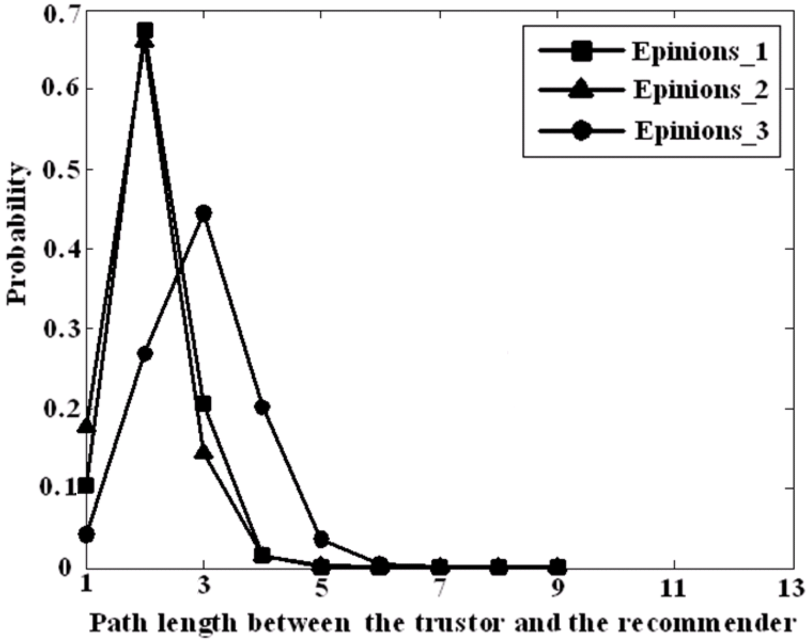

**Fig. 19.** Distribution of the path lengths from trustors to recommenders

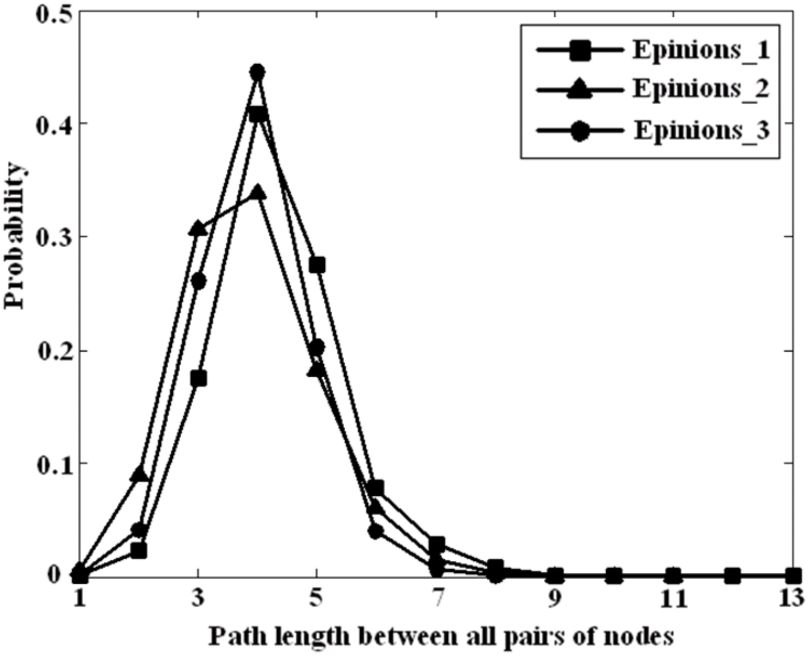

**Fig. 20.** Distribution of the path lengths between all pairs of nodes

Based on the analysis given in Chapter 3.3, the trust network has the scale-free structure, which is one structure of the small-world network. The structure of the scale-free network is shown in Fig. 21. According to the properties of the scale-free network, most nodes of the trust network have a few connections with other nodes, while a few nodes of the trust network have a large number of connections with other nodes, dominating the connectivity of the trust network. In addition, the trust network will continuously have the scale-free structure because of the continuous scale-freeness of the scale-free network. This contributes to the continuous effectiveness of using average path length of the trust network as MATD.

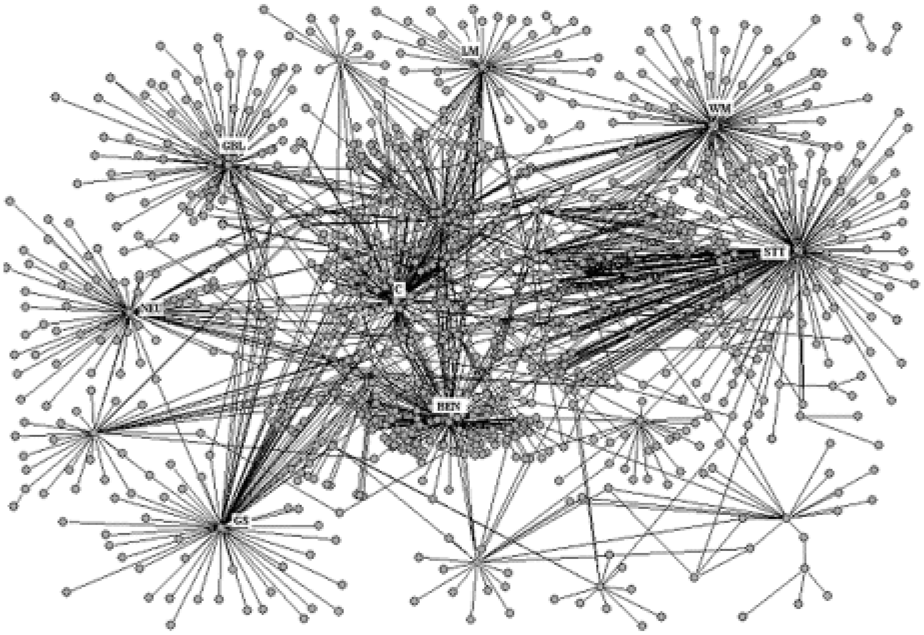

**Fig. 21.** Structure of the scale-free network

# Chapter 5

# Improved TARS using implicit trust networks

Existing works of TARS [1-7] focus on using the explicit trust. That is, the trust should be explicitly pointed out by the users. The limitation of this is that it is sometimes time consuming or expensive to get the explicit trust. This is because the explicit trust needs extra user efforts: users need to specifically point out their personal opinions on the trustees. What's more, in most practical recommender systems, these explicit trust statements are not available. This chapter proposes to improve the existing TARS model by using the implicit trust network: instead of using the effort-consuming explicit trust in TARS, other cheap and easy available trust sensitive information is used to generate the implicit trust for TARS. In particular, this work generates the implicit trust based on the user similarity. By comparing two users' ratings on their co-rated items, it is easy to get their similarity, as did in the conventional CF. This does not need extra human efforts on labeling the trust statements. The implicit trust is propagated among users and the implicit trust network is therefore constructed for TARS to achieve higher rating prediction coverage.

## 5.1 Building implicit trust networks for TARS

Since the explicit trust is not always available in the practical recommender systems, this work improves the conventional TARS by using the implicit trust network. The trust statement is regarded as implicit if it is not explicitly pointed out by the users. This work generates the implicit trust based on the user similarity.

Using the user's ratings on the items, the following steps are applied to build the implicit trust networks for TARS:

(1) Calculate the user similarity between users. The Pearson correlation coefficient, which is one of the most successful mechanisms in terms of accuracy in the conventional CF, is used to measure the user similarity in this research. The user similarity between two randomly selected users $u_1$ and $u_2$ is calculated as:

$$s_{u_1,u_2} = \frac{\sum_{I_{u_1,u_2}} (r_{u_1,i} - \overline{r_{u_1}})(r_{u_2,i} - \overline{r_{u_2}})}{\sqrt{\sum_{I_{u_1,u_2}} (r_{u_1,i} - \overline{r_{u_1}})^2 \sum_{I_{u_1,u_2}} (r_{u_2,i} - \overline{r_{u_2}})^2}},$$

in which $r_{u_1,i}$ and $r_{u_2,i}$ are $u_1$ 's rating and $u_2$ 's rating on item *i* respectively, $\overline{r_{u_1}}$ and $\overline{r_{u_2}}$ are $u_1$ 's average rating and $u_2$ 's average rating on all their rated items respectively, and $I_{u_1,u_2}$ is the items that are rated by $u_1$ and $u_2$ simultaneously, i.e., their co-rated items. $s_{u_1,u_2} \in [-1,1]$, in which a positive value implies a positive association (the larger $s_{u_1,u_2}$ is, the more similar $u_1$ and $u_2$ are) and a negative value implies a negative association (the larger $s_{u_1,u_2}$ is, the less similar $u_1$ and $u_2$ are).

(2) Generate the implicit trust based on the user similarity. In case two users are positively associated, the higher similarity value they have, the more likely they would find the recommendations given by the other one be useful. So it is more likely for them to trust each other. This work therefore regards the users implicitly trust those who have high user similarities with them, and implicitly distrust those who have low user similarities with them. This situation is inverse in case the users are negatively associated. Since the negative user similarities greatly increase the computational complexity in the implicit trust propagations, this work only discusses the implicit trust with respect to the positive user similarity in this research.

The implicit trust would be measured in various ways. This work uses the binary measurement due to its simplicity and popularity [24]. Specifically, this work codes 1 if the trustor trusts the trustee and code 0 in other cases. The trust function between the two randomly selected users is represented as:

$$t_{u_1,u_2} = f(s_{u_1,u_2}, size(I_{u_1,u_2})) = \begin{cases} 1 & s_{u_1,u_2} > Thres_s \wedge size(I_{u_1,u_2}) > Thres_I \\ 0 & else \end{cases},$$

in which $f(.)$ is the mapping function from the user similarity to the implicit trust. In addition to the user similarity, another attribute $size(I_{u_1,u_2})$, i.e., the size of the users' co-rated items, is involved in $f(.)$. This attribute is added to ensure the statistical effectiveness of the user similarity. For instance, if two users only have one or two co-rated items, their user similarity is not enough to reflect their real relationship, so it is meaningless to generate the implicit trust between these users. $Thres_s$ is the threshold of the user similarity, and $Thres_I$ is the threshold of the number of the co-rated items. Since the user similarity is mutual, the implicit trust in this research is nondirectional, i.e., $t_{u_1,u_2} = t_{u_2,u_1}$.

(3) Build the implicit trust network based on the implicit trust statements. It is far from enough to use the implicit trust directly in TARS. This is because the trust matrix is always very sparse due to the sparseness of the user similarities: it is only possible for the users to have similarities with a few users since it is not realistic for the users to rate all the items. Taking the advantage of the transitivity of trust, the implicit trust network is built for TARS to achieve higher rating prediction coverage. The implicit trust network is the trust network constructed on the basis of the implicit trust: the users act as the nodes and their implicit trusts act as the edges. In this case, the users can build up the trust relationships between each other even they do not have the direct implicit trust.

An example is given in Fig.22 to illustrate how this work builds the implicit trust network. Ten users are involved in the recommender system, as shown in the left side of the figure. This work generates the implicit trust by setting $Thres_s$ =0.75 and $Thres_I$ =2. Six users have implicit trusts with others, i.e.,

$A,B,C,D,E,F$ . These six users act as the nodes and their trusts act as edges, an implicit trust network is therefore constructed, as shown in the right side of the figure. Due to the binary trust measurement used in this research, the graph used to represent the implicit trust network is the binary graph. That is, an edge between two users means these two users are mutually trusted, having the trust value 1, while no edge represents the absence of trust.

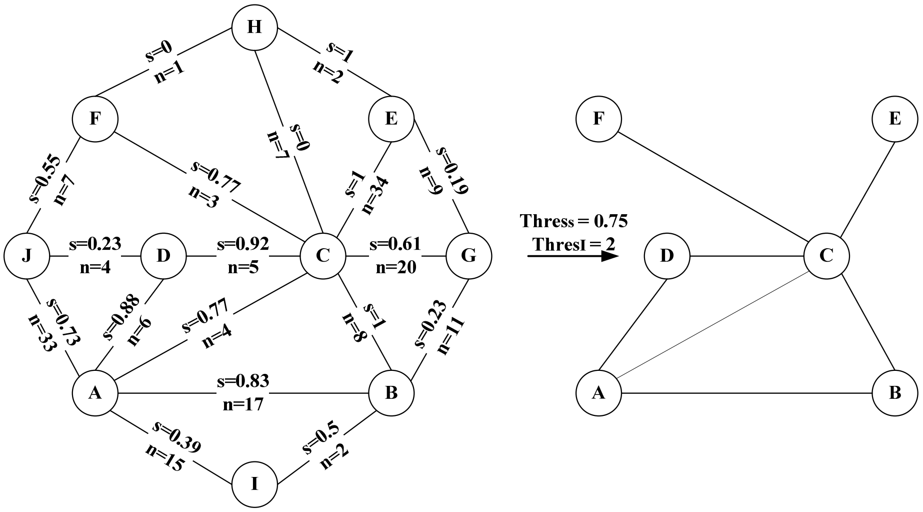

**Fig. 22.** An example of the implicit trust network generated by the user similarity

## 5.2 Finding small-world properties in implicit trust networks

The small-worldness of the trust network has been verified in chapter 3. The trust networks used for the experimental verification are the explicit trust networks, i.e., the trusts between all users of the trust network are explicitly pointed out by the users themselves. This subchapter verifies that the implicit trust networks, which are generated by the cheap or less effort-consuming trust related information, also have the small-world nature.

To verify the small-worldness of the implicit trust network, this work uses two kinds of verification methodology. On one hand, by using the conventional verification methodology, this work verifies that the implicit trust network has lager clustering coefficient and short average path length. On the other hand, this work verifies that the implicit trust network is the scale-free network. This indicates its continuous small-worldness in dynamic natures.

### 5.2.1 Implicit trust networks used in this research

Using the method shown in subchapter 5.1, three implicit trust networks are extracted from the Epinions dataset[9] to verify the small-worldness of the trust network. The Epinions dataset has two kinds of files: the rating data and the trust data. The rating data records the users' ratings on items. The trust data records the users' trust on other users. These trust statements are explicitly pointed out by the users. This work chooses to use the Epinions dataset to facilitate further comparisons between the explicit trust based TARS and the implicit trust based TARS. For the experiments held in this subchapter, only the rating data of the Epinions dataset are used.

Firstly, this work extracted three rating matrices from the Epinions dataset based on the timestamp of the ratings. They are named as $\mathbf{R}_{Epinions\_1}$, $\mathbf{R}_{Epinions\_2}$ and $\mathbf{R}_{Epinions\_3}$ respectively. $\mathbf{R}_{Epinions\_1}$ records the ratings stated in January 2001. It consists of 31019 users' 8632163 ratings on 551392 items. $\mathbf{R}_{Epinions\_2}$ records the ratings stated in the year 2002, from January to December. It consists of 2275 users' 740422 ratings on 36144 items. $\mathbf{R}_{Epinions\_3}$ records the ratings stated in November and December of 2003. It consists of 20157 users' 664061 ratings

[9] http://www.trustlet.org/wiki/Epinions_dataset

on 139633 items. $\mathbf{R}_{Epinions\_1}$ and $\mathbf{R}_{Epinions\_2}$ are extracted from the "extended epinions dataset"[10]. $\mathbf{R}_{Epinions\_3}$ is extracted from the "epinions dataset"[11].

Secondly, three implicit trust networks are constructed based on the above rating matrices. Specifically, this work sets $Thres_s$ =0.75 and $Thres_I$ =2 for the implicit trust generation function. The values of the thresholds are chosen based on the analysis of the explored rating matrices. $\mathbf{R}_{Epinions\_3}$ is used as an example to illustrate this. In $\mathbf{R}_{Epinions\_3}$, 19859 users have at least one co-rated item with other users, and there are totally 11160113 pairs of user similarities between these users. However, majority user similarities are useless for the implicit trust generation, as shown in Fig. 23: the user similarity of 90.08% pairs of users equals to 0, which means these users do not have any similarity. This is because most pairs of users only have limited number of co-rated items, as shown in Fig. 24: in case the user similarity equals to 0, 98.63% pairs only have one or two co-rated items. So this work only focuses on the pairs of users that have at least 3 co-rated items, i.e., $Thres_I$ =2 for the implicit trust generation. This would greatly reduce the useless information and make the implicit trust generation process more efficient. What's more, since a Pearson correlation greater than 0.7 is regarded as the strong positive association [22], this work sets $Thres_s$ =0.75 in this research.

[10] http://www.trustlet.org/wiki/Extended_Epinions_dataset.

[11] http://www.trustlet.org/wiki/Downloaded_Epinions_dataset.

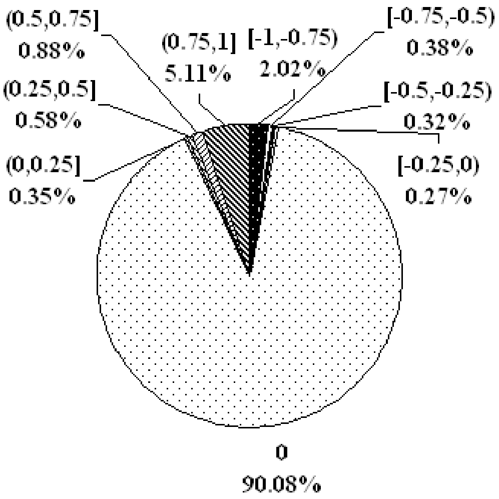

**Fig. 23.** Distribution of the user similarities between users

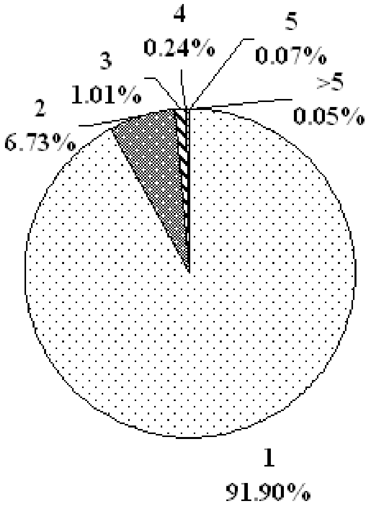

**Fig. 24.** Distribution of the co-rated items between users given user similarity equals to 0

The trust matrices generated from $\mathbf{R}_{Epinions\_1}$ , $\mathbf{R}_{Epinions\_2}$ and $\mathbf{R}_{Epinions\_3}$ are named as $\mathbf{T}_{Epinions\_1}$ , $\mathbf{T}_{Epinions\_2}$ and $\mathbf{T}_{Epinions\_3}$ respectively. $\mathbf{T}_{Epinions\_1}$ records 13244 users' 521507 implicit trust statements on each other. $\mathbf{T}_{Epinions\_2}$ records 1260 users' 54643 implicit trust statements on each other. $\mathbf{T}_{Epinions\_3}$ records 14704 users' 355033 implicit trust statements on each other. The implicit trust networks constructed by these three trust matrices are named as Epinions_1, Epinions_2 and Epinions_3 respectively. The characteristics of these implicit trust networks are summarized in Table 13. All users involved in these trust networks act as the trustors, the trustees or both.

**Table 13** Description of the implicit trust networks used in this research

| | **Number of nodes** | **Number of edges per node** |
|---|---|---|
| **Epinions_1** | 13244 | 39.38 |
| **Epinions_2** | 1260 | 43.38 |
| **Epinions_3** | 14704 | 24.15 |

### 5.2.2 Small-world characteristics of implicit trust networks

By analyzing Epinions_1, Epinions_2 and Epinions_3, this work verifies that the implicit trust network is the small-world network.

Firstly, the conventional experimental verification method is used to show the small-worldness of the implicit trust network. Similar as the explicit trust networks, the implicit trust networks also have large clustering coefficients. This is verified by comparing the clustering coefficients of the implicit trust networks and those of their corresponding random networks: using equation (3) and equation (4), the clustering coefficients of the explored implicit trust networks and their corresponding random networks are evluated, which are summarized in Table 14. The detailed distributions of the explored implicit trust networks' clustering coefficients are given in Fig. 25. It shows that similar as the explicit trust networks, though the clustering coefficients of some users are small (near 0), those of the majority users are greater than 0.1. A portion of the clustering coefficients even equals to 1. This means the neighbors of some users are fully connected. This is very different from the random network. The comparison between the clustering coefficients of the implicit trust networks and those of their corresponding random networks clearly shows that: the implicit trust network has much larger (higher order of magnitude) clustering coefficients than its corresponding random network. This satisfies the first condition of the small-world network's definition.

**Table 14** Clustering coefficients of the implicit trust networks and their corresponding random networks

| | $n$ | $k$ | $C$ | $C^R$ |
|---|---|---|---|---|
| **Epinions_1** | 13244 | 39.38 | 0.13 | $2.97\times10^{-3}$ |
| **Epinions_2** | 1260 | 43.38 | 0.62 | $3.44\times10^{-2}$ |
| **Epinions_3** | 14704 | 24.15 | 0.16 | $1.64\times10^{-3}$ |

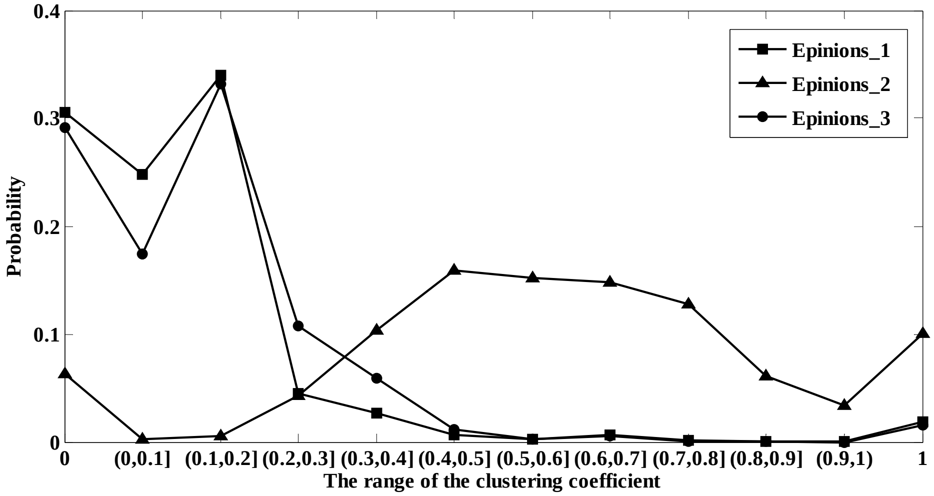

**Fig. 25.** Distribution of the implicit trust networks' clustering coefficients

Similar as the explicit trust networks, the implicit trust networks also have short average path lengths. This work chooses around 5% random samples from the large implicit trust networks (Epinions_1 and Epinions_3) and all samples from the small networks (Epinions_2) to analyze the average path length. The distributions of the explored implicit trust networks' average path lengths are given in Fig. 26. It shows that the implicit trust networks have very small number of direct trusts, i.e., where the path length equals to 1(less than 10% for Epinions_2, less than 1% for Epinions_1 and Epinions_3). By propagating trust, users can build up their trust relationships with others within several hops. Another important observation is that very small number of the trust propagations has long distance, e.g. the probabilities that the path lengths are longer than 5 hops (if any) are less than 5%. The path length of most trust propagations is from 2 hops to 4 hops. In more details: (1) the maximum path length of Epinions_1 is 7 hops, and its average path length is 2.73 hops; (2) the maximum path length of Epinions_2 is 4 hops, and its average path length is 2.05 hops; (3) the maximum path length of Epinions_3 is 10 hops, and its average path length is 2.98 hops.

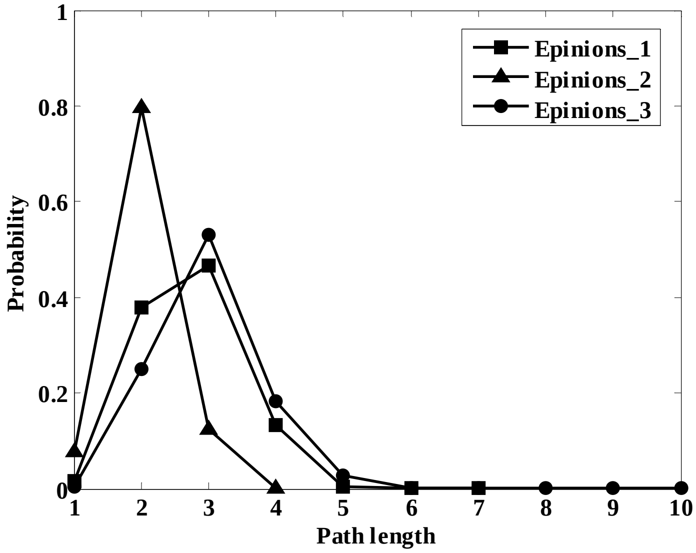

**Fig. 26.** Distribution of the implicit trust networks' path lengths

Using equation (5), the average path lengths of the explored implicit trust networks' corresponding random networks are evaluated, which are summarized in Table 15. Comparing the average path lengths of the implicit trust networks with those of their corresponding random networks, it is obvious that the implicit trust networks have similar (the same order of magnitude) average path lengths as their corresponding random networks. This satisfies the second condition of the small-world network's definition.

**Table 15** Average path lengths of the implicit trust networks and their corresponding random networks

| | $n$ | $k$ | $L$ | $L^R$ |
|---|---|---|---|---|
| **Epinions_1** | 13244 | 39.38 | 2.73 | 2.52 |
| **Epinions_2** | 1260 | 43.38 | 2.05 | 1.89 |
| **Epinions_3** | 14704 | 24.15 | 2.98 | 2.55 |

Using the above characteristics on the clustering coefficient and the average path length, this work compares the implicit trust networks with some well-known small-world networks documented in the literatures. These well-known small-world networks are those shown in Table 6. A further comparison between the small-world characteristics of the implicit trust networks and these networks is presented in Fig. 27. The axes of Fig. 27 represent the ratios of the selected networks and their corresponding random networks. Note that most small-world networks are concentrated around where the average path length ratio equals to 1. This means that the selected networks have similar average path length as their corresponding random networks. In addition, the clustering coefficient ratios of most networks are greater than 10. This means that the selected networks have much larger clustering coefficients than their corresponding random networks. The comparisons of Fig. 27 clearly show that the implicit trust networks have the same properties as other well-known small-world networks: they are highly clustered yet have small average path lengths. This work therefore draws the conclusion that the implicit trust networks are the small-world networks.

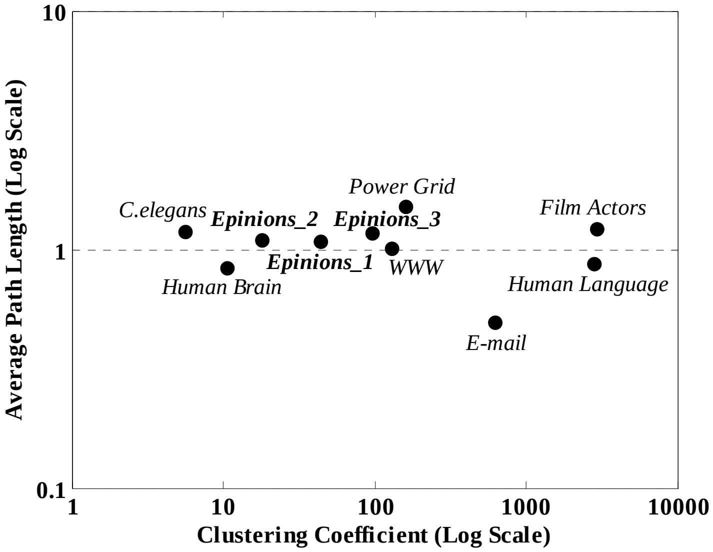

**Fig. 27.** Small-world characteristics of the implicit trust networks and some well-known small-world networks

As described in subchapter 3.3, the conventional experimental verification method only verifies the small-worldness of the implicit trust networks in static state. This work further verifies that the dynamically changing implicit trust networks still have the small-world characteristics. To achieve this, this work verifies that the implicit trust network is the scale-free network, as analyzed in subchapter 3.3

The degree distributions of the explored three implicit trust networks are calculated. The results are shown in Fig. 28, Fig. 29 and Fig. 30. Different as the explicit trust networks, this work does not differentiate the indegree distribution and the outdegree distribution of the implicit trust networks. This is because the explicit trust network is the directed network, while the implicit trust network used in this research is the undirected network: the implicit trusts are generated from the user similarity; since the user similarity between the users are bidirectional, the implicit trusts are bidirectional. It is clearly shown in Fig. 28, Fig. 29 and Fig. 30.that the nodes' degree distribution follows the power-law: $P(k) \sim k^{-\gamma}$, in which $P(k)$ is the probability that a randomly selected node has exactly $k$ edges, and $\gamma$ represents the power of the degree distribution. The detailed information about the degree distributions is shown in Table 16.

Since the implicit trust network is scale-free, based on the deduction shown in subchapter 3.3, this work draws the conclusion that the implicit trust networks are the small-world networks despite of their dynamics.

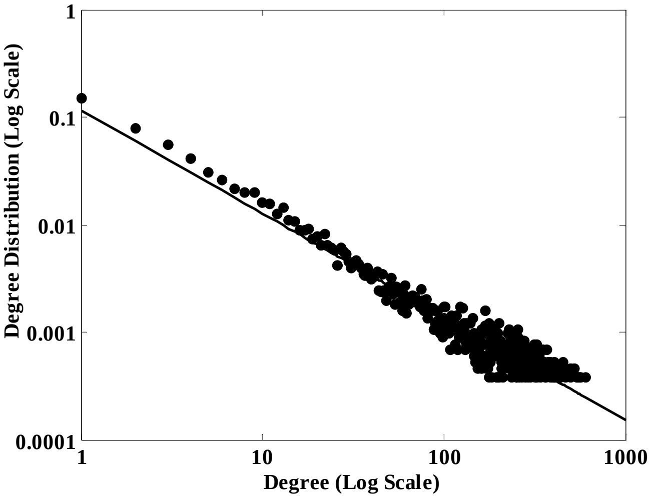

**Fig. 28.** Degree distribution of Epinions_1

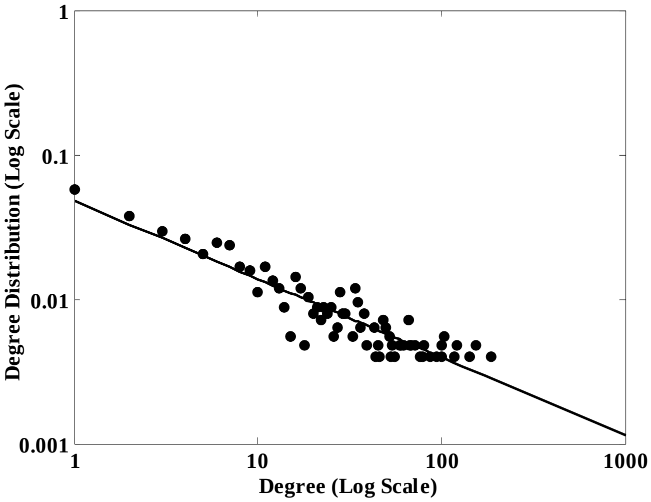

**Fig. 29.** Degree distribution of Epinions_2

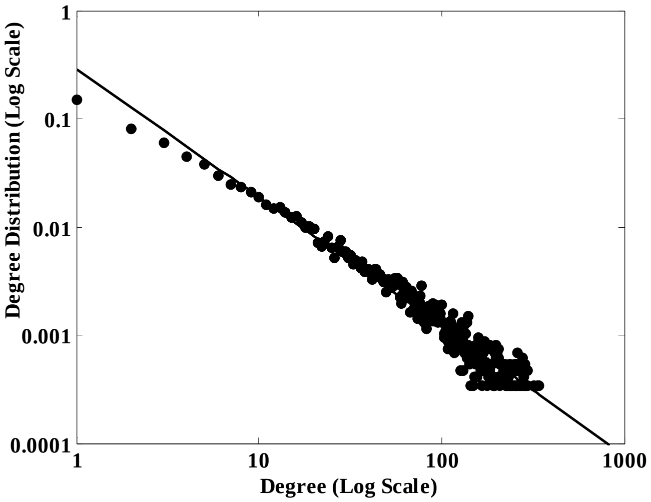

**Fig. 30.** Degree distribution of Epinions_3

**Table 16** Degree distributions of the implicit trust networks

| | $n$ | $k$ | $\gamma$ |
|---|---|---|---|
| **Epinions_1** | 13244 | 39.38 | 0.96 |
| **Epinions_2** | 1260 | 43.38 | 0.54 |
| **Epinions_3** | 14704 | 24.15 | 1.19 |

## 5.3 TARS using the small-worldness of implicit trust networks

Since the explicit trusts used in the conventional TARS models are not always available, this work proposes a novel TARS model using the implicit trust networks to improve the conventional ones. The proposed model is based on the verified small-worldness of the implicit trust networks.

### 5.3.1 The proposed TARS model

The architecture of the proposed TARS model is presented in Fig. 31. This model is based on the small-worldness of the implicit trust network. The input is the rating matrix which represents the ratings given by users on the items. The output is the matrix of the predicted ratings that the users would assign to the items. The black boxes in Fig. 31 represent various modules and the white boxes represent the matrices. The dash lines and dash boxes are used to show the architectures of the conventional CF [23] and the conventional TARS [3] in Fig. 31. The proposed method differs from the conventional CF in that the user similarity is further transformed to the implicit trust, and rating prediction is based on the implicit trust network. The proposed method differs from the conventional TARS in that the explicit trust is not needed in the rating prediction, while the implicit trust generated from the user similarity is used together with the rating matrix to predict the ratings.

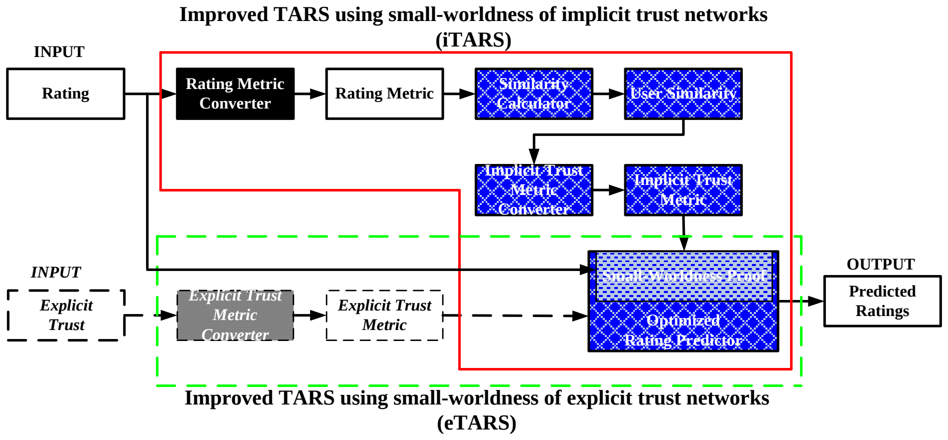
Improved TARS using small-worldness of implicit trust networks
(iTARS)
INPUT
Rating
Rating Metric Converter
Rating Metric
Similarity Calculator
User Similarity
Implicit Trust Metric Converter
Implicit Trust Metric
INPUT
Explicit Trust
Explicit Trust Metric Converter
Explicit Trust Metric
Small-Worldness Proof
Optimized Rating Predictor
OUTPUT
Predicted Ratings
Improved TARS using small-worldness of explicit trust networks
(eTARS)

**Fig. 31.** Architecture of the proposed implicit trust network based TARS model

The architecture has three modules: the similarity metric module, the implicit trust metric module and the rating predictor module. The similarity metric module is used to evaluate the user similarities between all users of the rating matrix. The implicit trust metric module is used to generate the implicit trust based on the user similarities. The details of these two modules have been discussed in subchapter 5.1. The rating predictor module is used to predict the ratings based on the recommendations given by various recommenders. In the module this work uses the rating prediction algorithm used shown in Table 8. That is, this work uses the same rating prediction mechanism as the one used in the proposed explicit trust network based TARS. The difference is that: in this model, the implicit trusts are generated from the user similarities, while the proposed  model shown in subchapter 4.1 requires the users to explicitly point out their trust on others.

### 5.3.2 Experimental results

The performances of TARS are examined to show the effectiveness of the proposed method. The experiments are held on the data shown in subchapter

5.2.1. $\mathbf{R}_{Epinions\_1}$, $\mathbf{R}_{Epinions\_2}$ and $\mathbf{R}_{Epinions\_3}$ are used as the inputs of the proposed method respectively. Predicted ratings on the items of these three rating matrices act as the outputs. $\mathbf{T}_{Epinions\_1}$, $\mathbf{T}_{Epinions\_2}$ and $\mathbf{T}_{Epinions\_3}$ records the implicit trust in the rating prediction procedure. Since the scales of $\mathbf{R}_{Epinions\_1}$, $\mathbf{R}_{Epinions\_2}$ and $\mathbf{R}_{Epinions\_3}$ are very large, it is computational expensive to predict ratings on all the items for all the users. This work chooses around 5% random samples from $\mathbf{R}_{Epinions\_1}$, $\mathbf{R}_{Epinions\_2}$ and $\mathbf{R}_{Epinions\_3}$ as the object of the rating prediction, i.e., this work predicts around 400000 ratings for $\mathbf{R}_{Epinions\_1}$, around 40000 ratings for $\mathbf{R}_{Epinions\_2}$ and around 30000 ratings for $\mathbf{R}_{Epinions\_3}$.

Firstly, this work verifies that $\lceil L \rceil$ is an appropriate value of MAPD for the implicit trust network based TARS, as claimed in the first phase of the rating predictor module. Using the implicit trust networks Epinions_1, Epinions_2 and Epinions_3, ratings are predicted on the rated items of $\mathbf{R}_{Epinions\_1}$, $\mathbf{R}_{Epinions\_2}$ and $\mathbf{R}_{Epinions\_3}$ respectively. The MAPD of the proposed method is calculated by equation (6): for Epinions_1, $d_{\max} = \lceil L \rceil \approx \lceil 2.52 \rceil = 3$; for Epinions_2, $d_{\max} = \lceil L \rceil \approx \lceil 1.89 \rceil = 2$; for Epinions_3, $d_{\max} = \lceil L \rceil \approx \lceil 2.55 \rceil = 3$.

The rating prediction coverage of TARS is examined to verify the effectiveness of MAPD in the proposed method. The coverage of TARS is measured by both the rating coverage and the recommender coverage. The rating coverage and the recommender coverage of the proposed model by using different values of MAPD are reported in Table 17 and Table 18 respectively, in which the bold values are the coverage calculated by using the suggested MAPD values. The active users can build up their implicit trust relationships with the recommenders within 5 hops in $\mathbf{R}_{Epinions\_1}$, within 4 hops in $\mathbf{R}_{Epinions\_2}$ and within 7 hops in $\mathbf{R}_{Epinions\_3}$. The experimental results show that: (1) If the value of MAPD is set to be smaller than the suggested value, both the rating coverage and the recommender coverage of TARS decrease, in which the recommender coverage decreases significantly. (2) If the value of MAPD is set to be greater than the suggested value, the rating coverage and the recommender coverage of TARS do not change significantly. This work therefore draws the conclusion that $\lceil L \rceil$ is a suitable value of MAPD for the proposed model of TARS.

**Table 17** Recommender coverage of the proposed method by using different values of MAPD

| | $\mathbf{R}_{Epinions\_1}$ | $\mathbf{R}_{Epinions\_2}$ | $\mathbf{R}_{Epinions\_3}$ |
|---|---|---|---|
| $d_{\max}$ **= 1** | 23.04% | 42.71% | 6.81% |
| $d_{\max}$ **= 2** | 95.50% | **95.93%** | 65.52% |
| $d_{\max}$ **= 3** | **98.11%** | 96.66% | **85.12%** |
| $d_{\max}$ **= 4** | 98.20% | 96.66% | 86.16% |
| $d_{\max}$ **= 5** | 98.20% | - | 86.21% |
| $d_{\max}$ **= 6** | - | - | 86.21% |
| $d_{\max}$ **= 7** | - | - | 86.21% |

**Table 18** Rating coverage of the proposed method by using different values of MAPD

| | $\mathbf{R}_{Epinions\_1}$ | $\mathbf{R}_{Epinions\_2}$ | $\mathbf{R}_{Epinions\_3}$ |
|---|---|---|---|
| $d_{\max}$ **= 1** | 92.30% | 98.10% | 47.25% |
| $d_{\max}$ **= 2** | 99.53% | **99.10%** | 82.44% |
| $d_{\max}$ **= 3** | **99.53%** | 99.10% | **91.17%** |

| | | | |
|---|---|---|---|
| $d_{\max}$ **= 4** | 99.53% | 99.10% | 92.19% |
| $d_{\max}$ **= 5** | 99.53% | - | 92.40% |
| $d_{\max}$ **= 6** | - | - | 92.40% |
| $d_{\max}$ **= 7** | - | - | 92.40% |

Note that $\lceil L \rceil$ is only similar to the average trust propagation distance between two randomly selected users of the implicit trust network, but the experiments show that $\lceil L \rceil$ is a appropriate value of MAPD for the proposed TARS. This is because it is the average trust propagation distance between all pairs of users that $\lceil L \rceil$ is similar to. However, not all users are recommenders. Further analysis on the distribution of the average path length between the active users and the recommenders, which is given in Fig. 32, shows that: compared with the distribution of the average path length between all pairs of users in the implicit trust network, as shown in Fig. 26, the average path length between the active users and recommenders are much smaller than that between all pairs of users, and the maximum distance between the active users and the recommenders are always shorter than that between all pairs of users. This indicates that compared with the non-recommenders or the non-active recommenders, the recommenders tend to have shorter distances with the active users. This contributes to the effectiveness of the method by setting $\lceil L \rceil$ as the value of MAPD for TARS. This phenomenon is the same as the proposed TARS model in Chapter 4.

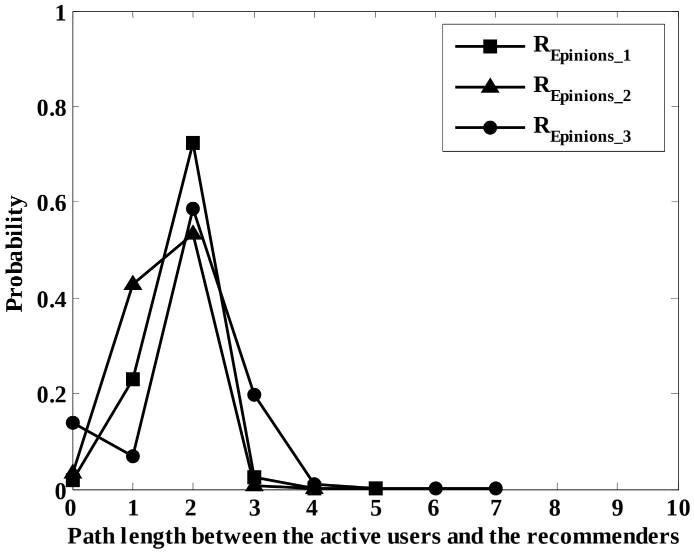

**Fig. 32.** Distribution of the path lengths between the active users and the recommenders

Secondly, the proposed model is compared with the proposed TARS model shown in subchapter 4.1 and the conventional CF. The performances of these models are examined in two aspects with the proposed model: the rating prediction accuracy and the rating prediction coverage (including the recommender coverage and the rating coverage). The rating prediction accuracy of the recommender system is measured by the error of the predicted ratings. Specifically, this work calculates the Mean Absolute Error (MAE). By predicting the rating on the rated items of $\mathbf{R}_{Epinions\_1}$, $\mathbf{R}_{Epinions\_2}$ and $\mathbf{R}_{Epinions\_3}$, the MAE of different models is reported in Table 19, in which iTARS represents the implicit trust network based TARS model proposed in this chapter, and eTARS represents the explicit trust network based TARS model proposed in Chapter 4. The recommender coverage and the rating coverage of different models are given in Table 20 ad Table 21 respectively. These experimental results show that:

(1) Comparing with eTARS: in contrast to the decreasing of the rating prediction coverage, the rating prediction accuracy is improved by iTARS.

Specifically, in the experiments held on $\mathbf{R}_{Epinions\_1}$, by decreasing 1.05% recommender coverage and 0.38% rating coverage of eTARS, iTARS can improve 19.44% of its rating accuracy; in the experiments held on $\mathbf{R}_{Epinions\_2}$, by decreasing 0.61% recommender coverage and 0.84% rating coverage of eTARS, iTARS can improve 12.00% of its rating accuracy; in the experiments held on $\mathbf{R}_{Epinions\_3}$, by decreasing 11.22% recommender coverage and 8.59% rating coverage of eTARS, iTARS can improve 34.86% of its rating accuracy.

(2) Comparing with the conventional CF: iTARS has similar rating prediction accuracy as the conventional CF, while the recommender coverage and the rating coverage are improved, in which the recommender coverage tends to be greatly improved. Specifically, in the experiments held on $\mathbf{R}_{Epinions\_1}$, the recommender coverage is 18.65% improved and rating coverage is 0.55% increased by using iTARS; in the experiments held on $\mathbf{R}_{Epinions\_2}$, the recommender coverage is 36.18% improved and rating coverage is 0.22% increased by using iTARS; in the experiments held on $\mathbf{R}_{Epinions\_3}$, the recommender coverage is 159.67% improved and rating coverage is 16.72% increased by using iTARS.

To sum up, the proposed implicit trust network based TARS model is superior to the conventional TARS not only in that it releases the user efforts in trust labeling, but also in that the proposed model improves the rating prediction accuracy with little cost in the rating prediction coverage; the proposed implicit trust network based TARS model is superior to the conventional CF in that the proposed model improves the rating prediction coverage without cost in the rating prediction accuracy. This work therefore draws the conclusion that the proposed model of TARS, which is based on the small-worldness of the implicit trust network, is effective in rating predictions.

**Table 19** MAE of the proposed model and the conventional models

| | $\mathbf{R}_{Epinions\_1}$ | $\mathbf{R}_{Epinions\_2}$ | $\mathbf{R}_{Epinions\_3}$ |
|---|---|---|---|
| **The proposed iTARS** | 0.29 | 0.22 | 0.71 |
| **The proposed eTARS** | 0.25 | 0.21 | 0.73 |
| **Conventional CF** | 0.29 | 0.20 | 0.73 |

**Table 20** Recommender coverage of the proposed model and the conventional models

| | $\mathbf{R}_{Epinions\_1}$ | $\mathbf{R}_{Epinions\_2}$ | $\mathbf{R}_{Epinions\_3}$ |
|---|---|---|---|
| **The proposed iTARS** | 98.11% | 95.93% | 85.12% |
| **The proposed eTARS** | 99.15% | 96.52% | 95.88% |
| **Conventional CF** | 82.69% | 70.44% | 32.78% |

**Table 21** Rating coverage of the proposed model and the conventional models

| | $\mathbf{R}_{Epinions\_1}$ | $\mathbf{R}_{Epinions\_2}$ | $\mathbf{R}_{Epinions\_3}$ |
|---|---|---|---|
| **The proposed iTARS** | 99.53% | 99.10% | 91.17% |
| **The proposed eTARS** | 99.91% | 99.94% | 99.74% |
| **Conventional CF** | 98.99% | 98.88% | 78.11% |

In the implicit trust network based TARS, the trust statements are generated by the trust sensitive information. In the explicit trust network based TARS, the trust statements are explicitly pointed out by the users. The above simulations results show that the proposed implicit trust network based TARS has similar rating prediction performance as the explicit trust network based TARS. Based on the analysis of the explicit trust and the generated implicit trust, it is found that there are some relationships between these two kinds of trust statements. This may contributes to the similar rating prediction accuracy of these two kinds of TARS models. The concrete relationships between the implicit trust statements and explicit trust statements used in the proposed TARS models are given in Fig. 33, Fig. 34 and Fig. 35. These simulation results clear show that: (1) if two users have short explicit trust propagation distance, they also tends to have short implicit trust propagation distance; (2) if two users have long explicit trust propagation distance, they also tends to have long implicit trust propagation distance.

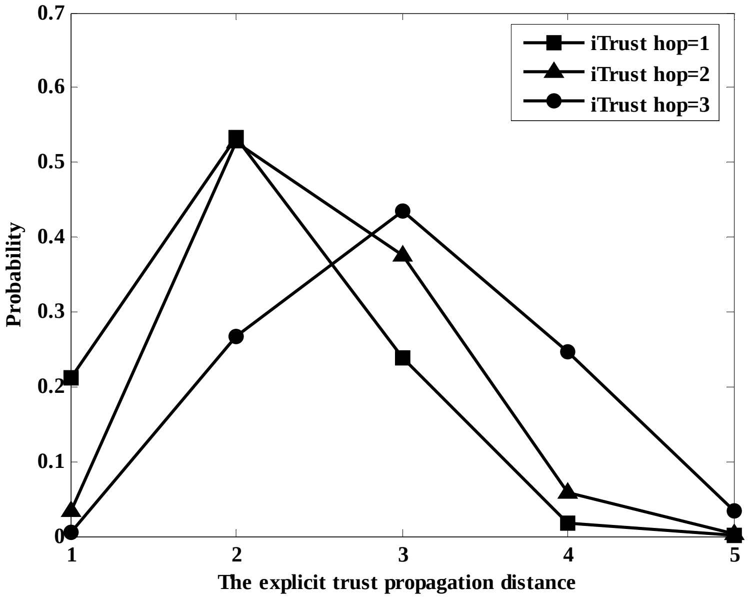

**Fig. 33.** Relationship of the implicit trust and the explicit trust used for predicting ratings of $\mathbf{R}_{Epinions\_1}$

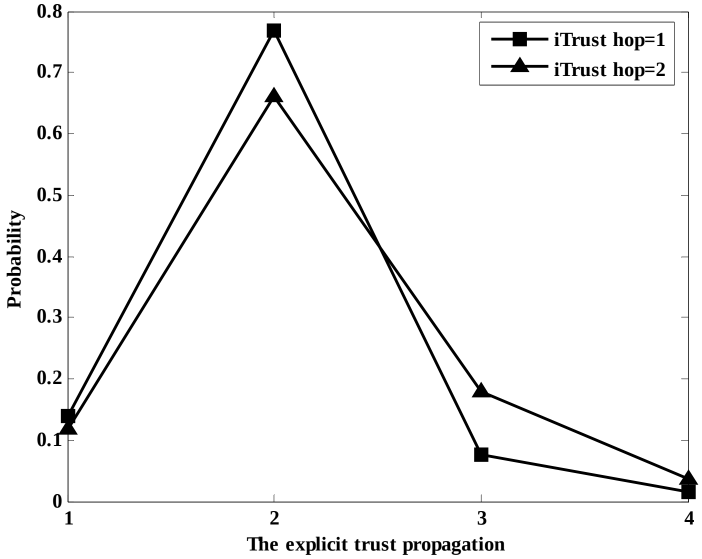

**Fig. 34.** Relationship of the implicit trust and the explicit trust used for predicting ratings of $\mathbf{R}_{Epinions\_2}$

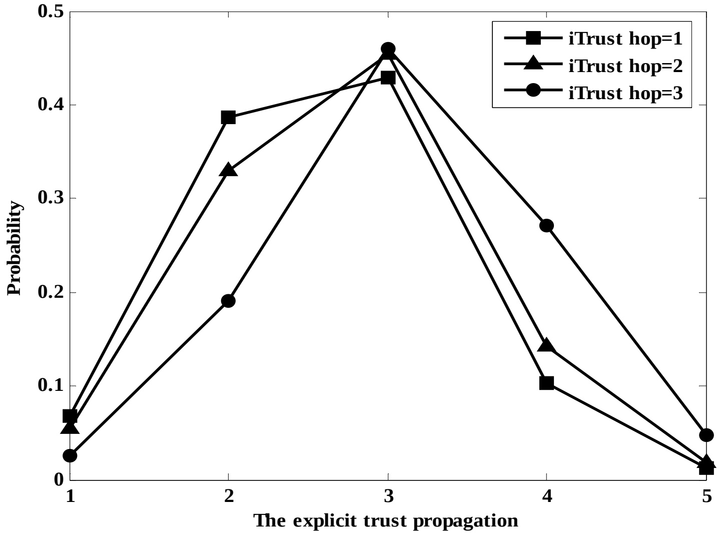

**Fig. 35.** Relationship of the implicit trust and the explicit trust used for predicting ratings of $\mathbf{R}_{Epinions\_3}$

# Chapter 6

# Conclusion and future works

Using the experimental data extracted from five public released datasets, this work verifies the small-worldness of the trust network: the nodes of trust network are highly clustered while the distance between two randomly select nodes is short. This work further verifies that the trust network continuously has the small-world structure. This is achieved by verifying the scale-freeness of the trust network. One basic property of the scale-free network is that its structure and dynamics are independent of its scale. This ensures the continuous scale-freeness of the scale-free network in dynamics. Since the scale-free network is one category of the small-world network, by verifying its scale-freeness, this work shows that the small-worldness of the trust network is independent of its dynamics.

The small-worldness of the trust network indicates that any two nodes of the trust network could be connected within limited number of trust propagations, and the average trust propagation distance is similar to the average path length of this trust network's corresponding random network, which is easy to calculate since it only relates to the size and the average degree of the trust network. This work uses this property to optimize the conventional trust-aware recommender system: the average path length of the trust network is used to approximately act as the value of the maximum allowable propagation distance of TARS. The performances of this optimized TARS model are examined on three large scale real data. The simulations results clearly show that the proposed optimized TARS model can achieve the maximum rating prediction accuracy and the maximum rating prediction coverage with the minimum computation complexity.

This thesis further propose a novel TARS model based on the small-worldness of the implicit trust network, in which the implicit trust is generated from the user similarities. Conventional TARS suffers from the problem that it needs extra user efforts to label the trust statements. The proposed model solves this problem by generating the implicit trust based on other "cheap" trust sensitive information, i.e., the information that needs little or no extra user efforts. In addition, the proposed model is able to improve the rating prediction accuracy of the conventional TARS with little cost in the rating prediction coverage. Conventional CF suffers from the data sparseness problem, that is, it is hard to find the user similarities between a number of active users and recommenders. Though the proposed model generates the implicit trust based on the sparse user similarities, this work solve the data sparseness problem by propagating the implicit trust and build the implicit trust network for the rating prediction. By analyzing the implicit trust networks of TARS, this work verifies the small-world topology of the implicit trust network. This indicates that, similar as the explicit trust network, the trust propagation distance between any two users of the implicit trust network is short, within limited number of trust propagation hops. Using the same rating prediction mechanism of the proposed optimized TARS model using the explicit trust network, experimental results show that the proposed trust-aware recommender system using the implicit trust network can also achieve high rating prediction accuracy and high rating prediction coverage.

The future work of this thesis will mainly focus on how to filter out the unfair recommendations for TARS. TARS suggests information to the active users based on the recommendations given by various recommenders. However, there may exist some self-interested recommenders who give unfair recommendations to maximize their own gains (perhaps at the cost of others). So it is essential to avoid or reduce the influence of the unfair positive or negative recommendations from the self-interested recommenders. For this purpose, I intend to introduce the users' distrust statements into the TARS model. By analyzing the recommendations given by each user's distrusted recommenders and the relationship between the trust statements and the distrust statements, the reliable recommendations will be chosen for the rating aggregations of the proposed TARS model.

# Bibliography

[1] Massa, P., Avesani, P.: Trust Metrics in Recommender Systems, Proceedings of Computing with Social Trust, 2009, pp. 259-285

[2] Massa, P., Avesani, P.: Trust-aware Collaborative Filtering for Recommender Systems, Proceedings of Federated Int. Conference on the Move to Meaningful Internet, 2004, pp. 492-508

[3] Li, Y., Kao, C., TREPPS: A Trust-based Recommender System for Peer Production Services, Expert Systems with Applications. 36 (2009) 3263-3277.

[4] Walter, F., Battiston, S., Schweitzer, F., A model of a trust-based recommendation system on a social network, Autonomous Agents and Multi-Agent Systems. 16 (2008) 57-74.

[5] Andersen, R., Borgs, C., Chayes, J., Feige, U., Flaxman, A., Kalai, A., et al., Trust-based recommendation systems: an axiomatic approach, in: Proceeding of the 17th International Conference on World Wide Web, Beijing, China, ACM, 2008: pp. 199-208.

[6] Coates, G., Hopkinson, K., Graham, S., Kurkowski, S.: Collaborative, Trust-Based Security Mechanisms for a Regional Utility Intranet, IEEE Transactions On Power System, 2008, 23, (3), pp. 831-844

[7] Massa, P., Avesani, P.: Controversial Users Demand Local Trust Metrics: An Experimental Study on Epinions.com Community, Proceedings of American Association for Artificial Intelligence 2005, 2005, pp.121-126

[8] Massa, P., Avesani, P.: Trust-aware recommender systems, Proceedings of the 2007 ACM Conference on Recommender Systems, 2007, pp. 17-24

[9] Musial, K., Recommender System for Online Social Network, Lap Lambert Academic Publishing, 2009.

[10] Hanneman, R., Riddle, M., Introduction to social network methods, Riverside, CA: University of California, Riverside. http://faculty.ucr.edu/~hanneman/, assessed November, 2009

[11] Castro, A., Vanhoof, S., Van, W., Onghena, P.: The Non-Transitivity of Pearson's Correlation Coefficient: An Educational Perspective, Proceedings of the 56th Session of the International Statistical Institute, 2007, pp. 22-29

[12] Newman, M., Barabasi, A., Watts, D.: The Structure and Dynamics of Networks (Princeton University Press, 1st ed., 2006)

[13] Missen, M.M.S., Boughanem, M., Gaume, B., The Small World of Web Network Graphs, in: Wireless Networks, Information Processing and Systems, 2009: pp. 133-145.

[14] Markosova, M., Nather, P.: Language as a Small World Network, Proceedings of the 6th International Conference on Hybrid Intelligent Systems, 2006, pp. 37-37

[15] Ebel, H., Mielsch, L., Bornholdt, S.: Scale-free topology of e-mail networks, Physical Review E, 66, 035103 (R), 2002

[16] Bullmore, E., Sporns, O., Complex brain networks: graph theoretical analysis of structural and functional systems, Nat. Rev. Neurosci. 10 (2009) 186-198.

[17] Bassett, D., Bullmore, E.: Small-World Brain Networks, Neuroscientist, 2006, 12, (6), pp. 512-523

[18] Amaral, L., Scala, A., Barthelemy M., Stanley H., Classes of small-world networks., Proceedings of the National Academy of Sciences of the United States of America. 97 (2000) 11152, 11149.

[19] Barabasi, A., Scale-Free Networks: A Decade and Beyond, Science. 325 (2009) 412-413.

[20] Watts, D., Strogatz, S.: Collective dynamics of 'small-world' networks, Nature, 1998, 393, pp. 440-442

[21] Artz D., Gil Y., A survey of trust in computer science and the Semantic Web, Web Semant. 5 (2007) 58-71.

[22] O’Donovan J., Smyth B., Trust in recommender systems, in: Proceedings of the 10th International Conference on Intelligent User Interfaces, San Diego, California, USA, ACM, 2005: pp. 167-174.

[23] Gray E., Seigneur J., Chen Y., C. Jensen, Trust propagation in small worlds, In Proc. Of 1st Int. Conf. on Trust Management (iTrust’03. (2003) 239—254.

[24] Adamic L.A., The Small World Web, in: Proceedings of the Third European Conference on Research and Advanced Technology for Digital Libraries, Springer-Verlag, 1999: pp. 443-452.

[25] Venkatraman M., Yu B., Singh M.P., Trust and reputation management in a small-world network, In Proceedings of Fourth International Conference on

MultiAgent Systems. (2000) 449—450.

[26] Guo X., Li X., Qin Y., Chen C., Modeling Small-World Trust Networks, in: Ubiquitous Multimedia Computing, 2008. UMC ’08. International Symposium On, 2008: pp. 154-159.

[27] Achard S., Salvador R., Whitcher B., Suckling J., Bullmore E., A Resilient, Low-Frequency,Small-World Human Brain Functional Network with Highly Connected Association Cortical Hubs, J.Neurosci. 26 (2006) 63-72.

[28] Newman M.E.J., Models of the Small World: A Review, Cond-Mat/0001118. (2000)

[29] Wang X.F., Chen G., “Complex networks: small-world, scale-free, and beyond,” IEEE Circ. Syst.Magazine, Vol. 3, No. 1, pp. 6-20, June 2003

[30] Barabasi A., Ravasz E., Vicsek T., Deterministic Scale-Free Networks, Cond-Mat/0107419. (2001)

[31] Massa P., Souren K., Trustlet, Open Research on Trust Metrics, in: Proceedings of the 2ndWorkshop on Social Aspects of the Web (SAW 2008), pp. 31-43.

[32] Bedi P. and Kaur H., Marwaha S., Trust Based Recommender System for Semantic Web, in: Proceedings of the 2007 International Joint Conferences on Artificial Intelligence, Hyderabad, India, pp. 2677-2682.

[33] Pitsilis G., Marshall L., A Trust-enabled P2P Recommender System, in: Proceedings of 15th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, 2006, pp. 59-64.

[34] Massa P., Avesani P., Controversial users demand local trust metrics: an experimental study on epinions.com community", in: Proceedings of 20th national conference on Artificial intelligence, 2005, pp. 121-126.

[35] Avesani P., Massa P., Tiella R., A trust-enhanced recommender system application: Moleskiing, in: Proceedings of the 2005 ACM symposium on Applied computing, pp. 1589 – 1593.

[36] Watts D. J., Small Worlds: The Dynamics of Networks Between Order and randomness, Princeton University Press, 1999.

[37] Ziegler, C.N., Golbeck, J. Investigating interactions of trust and interest similarity. Decision Support Systems, 2007, 43(2), 460–475.

[38] Zadeh, L. A. Fuzzy sets. Information and Control, (1965),8, 338–353.

[39] Wei, Y. Z., Moreau, L., Jennings, N. R. A market-based approach to recommender systems. ACM Transactions on Information Systems, 2005, .23(3), 227–266.

[40] Song, S., Hwang, K., Zhou, R., Kwok, Y.K. Trusted P2P transactions with fuzzy reputation aggregation. IEEE Internet Computing, 2005, 9(6), 24–34.

[41] Sinha, R., Swearingen, K. The role of transparency in recommender systems. Conference on human factors in computing systems. Minneapolis, MN, USA: ACM Press, pp. 830–831, 2002.

[42] Sinha, R., Swearingen, K. Comparing recommendations made by online systems and friends. In Proceedings of the DELOS-NSF workshop on personalization and recommender systems in digital libraries. Dublin, Ireland, 2001.

[43] Singh, M. P., Yu, B., Venkatraman, M. Community-based service location. Communications of the ACM, 2001, 44(4), 49–54.

[44] Sabater, J., & Sierra, C. Review on computational trust and reputation models. Artificial Intelligence Review, 2005, 24(1), 33–60.

[45] Sabater, J., Sierra, C. REGRET: A reputation model for gregarious societies. In Proceedings of the 4th workshop on deception fraud and trust in agent societies, 2001, pp. 61–70). Montreal, Canada.

[46] O'Donovan, J., Smyth, B. Trust in recommender systems. In Proceedings of the 10th international conference on intelligent user interfaces, 2005, pp. 167–174. San Diego, CA, USA: ACM Press.

[47] McKnight, D. H., Chervany, N. L. The meanings of trust. Technical Report WP9604. University of Minnesota Management Information Systems Research Center, 1996.

[48] Mamdani, E. H., Assilian, S. An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Human-Computer Studies, 1999, 51(2), 135–147.

[49] Liou, T.S., Wang, M.J. J. Fuzzy weighted average: An improved algorithm. Fuzzy Sets and Systems, 1992, 49(3), 307–315.

[50] Lindahl, C., Blount, E. Weblogs: Simplifying web publishing.Computer, 2003, 36(11), 114–116.

[51] Li, Y.M., Li, T.Y., Chen, J.C. Trust based instant messaging system. In Proceedings of the 9th conference on information management practice, 2006, Yunlin, Taiwan (in Chinese).

[52] Li, Y.M., Chen, J.C., Li, T.Y. A Blog system with trust mechanism. In Proceedings of the 18th international conference on information management, 2007, Taipei, Taiwan (in Chinese)

[53] Li, D.F. Compromise ratio method for fuzzy multi-attribute group decision making. Applied Soft Computing, 2007, 7(3), 807–817.

[54] Kuo, Y.L., Yeh, C.H., Chau, R. A validation procedure for fuzzy multiattribute decision making. The 12th IEEE International Conference on Fuzzy Systems, 2003, 2, 1080–1085.

[55] Kuo, M.S., Tzeng, G.H., Huang, W.C. Group decisionmaking based on concepts of ideal and anti-ideal points in a fuzzy environment. Mathematical and Computer Modelling, 2007, 45(3–4), 324–339.

[56] Kolbitsch, J., Maurer, H. The transformation of the web: How emerging communities shape the information we consume. Journal of Universal Computer Science, 2006, 12(2), 187–213.

[57] Kleinberg, J. Navigation in a small world. Nature, 2000, 406, 845.

[58] Joang, A. The right type of trust for distributed systems. In Proceedings of the 1996 workshop on new security paradigms, 1996, pp. 119–131. Lake Arrowhead, CA, United States: ACM Press.

[59] Joang, A., Ismail, R., Boyd, C. A survey of trust and reputation systems for online service provision. Decision Support Systems, 2007, 43(2), 618–644.

[60] Hwang, C.L., Yoon, K. Multiple attribute decision making: methods and applications: A state-of-the-art survey. Berlin: Springer-Verlag, 1981.

[61] Huynh, T. D., Jennings, N. R., Shadbolt, N. R. An integrated trust and reputation model for open multi-agent systems. Journal of Autonomous Agents and Multi-Agent Systems, 2006, 13(2), 119–154.

[62] Heath, T., Motta, E., Petre, M. Person to person trust factors in word of mouth recommendation. In Proceedings CHI2006 workshop on reinventing trust, collaboration, and compliance in social systems, 2006.

[63] Golbeck, J. Personalizing applications through integration of inferred trust values in semantic web-based social networks. In Proceedings of the semantic

network analysis workshop at the 4th international semantic web conference, 2005.

[64] Golbeck, J. Generating predictive movie recommendations from trust in social networks. In Proceedings of the 4th international conference on trust management. Pisa, Italy, 2006.

[65] Gantz, J. F., Reinsel, D., Chute, C., Schlichting, W., McArthur, J.,Minton, S., et al. The expanding digital universe: A forecast of worldwide information growth through 2010. EMC. IDC, 2007.

[66] Dubois, D., Prade, H. Fuzzy sets and systems: Theory and applications. NewYork: Academic Press, 1980.

[67] Ding, L., Zhou, L., Finin, T. Trust based knowledge outsourcing for semantic web agents. In Proceedings of the 2003 IEEE/WIC international conference on web intelligence, pp. 379–387.

[68] Ding, L., Kolari, P., Ganjugunte, S., Finin, T., Joshi, A. Modeling and evaluating trust network inference. In Proceedings of the 7th international workshop on trust in agent societies. New York, 2004.

[69] Chen, S.J., Hwang, C. L. Fuzzy multiple attribute decision making: Methods and applications. New York: Springer-Verlag, 1992.

[70] Chen, S.J. A new similarity measure of generalized fuzzy numbers based on geometric-mean averaging operator. IEEE International Conference on Fuzzy Systems. BC, Canada: Vancouver, 2006, pp. 1879–1886.

[71] Chen, M.F., Tzeng, G.H. Combining grey relation and TOPSIS concepts for selecting an expatriate host country. Mathematical and Computer Modelling, 2004, 40, 1473–1490.

[72] Chen, C.T. Extensions of the TOPSIS for group decision-making under fuzzy environment. Fuzzy Sets and Systems, 2000, 114, 1–9.

[73] Chang, P.T., Hung, K.C., Lin, K.P., Chang, C.H. A comparison of discrete algorithms for fuzzy weighted average. IEEE Transactions on Fuzzy Systems, 2006, 14(5), 663–678.

[74] Castellano, G., Fanelli, A. M., Mencar, C.. Design of transparent Mamdani fuzzy inference systems. Design and Application of Hybrid Intelligent Systems, 2003, 468–476.

[75] Benkler, Y. The wealth of networks: How social production transforms

markets and freedom. Yale University Press, 2006.

[76] Bellman, R. E., Zadeh, L. A. Decision-making in a fuzzy environment. Management Science, 1970, 17(4), B141–B164.

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

[78] Abdul-Rahman, A., Hailes, S. A distributed trust model. In Proceedings of the 1997 workshop on new security paradigms, 1998, pp. 48–60, Langdale, Cumbria, United Kingdom: ACM Press.

# List of publications

## Journals

[1] Weiwei Yuan, Donghai Guan, Young-Koo Lee, Sungyoung Lee and Sung Jin Hur, “Improved trust-aware recommender system using small-worldness of trust networks”, Knowledge-Based Systems 23 (2010) 232-238. (SCI)

[2] Weiwei Yuan, Donghai Guan, Young-Koo Lee and Sungyoung Lee “iTARS: Trust-Aware Recommender System using Implicit Trust Networks”, IET Communications, accepted. (SCI)

[3] Weiwei Yuan, Donghai Guan, Young-Koo Lee and Sungyoung Lee “The Small-World Trust Network”, Applied Intelligence, , accepted. (SCI)

[4] Donghai Guan, Weiwei Yuan, Young-Koo Lee, and Sungyoung lee. Nearest Neighbor Editing Aided by Unlabeled Data. Information Sciences, Vol 179, Issue 13, pp. 2273-2282, 2009. (SCI)

[5] Donghai Guan, Weiwei Yuan, Young-Koo Lee and Sungyoung Lee, "Identifying mislabeled training data with the aid of unlabeled data", Applied Intelligence, Accepted. (SCI)

[6] Donghai Guan, Weiwei Yuan, Young-Koo Lee, Andrey Gavrilov, and Sungyoung Lee. Improving Supervised Learning Performance by Using Fuzzy Clustering Method to Select Training Data. Journal of Intelligent & Fuzzy Systems, Vol 19, pp. 321-334, 2008. (SCIE)

[7] Weiwei Yuan, Donghai Guan, Sungyoung Lee, and Youngkoo Lee, "A Dynamic Trust Model Based on Naive Bayes Classifier for Ubiquitous Environments", The 2006 International Conference on High Performance Computing and Communications (HPCC-06). LNCS 4208, ISBN 3-540-39368-4, pp.562-571. (LNCS, SCIE)

[8] Weiwei Yuan, Donghai Guan, Sungyoung Lee and Youngkoo Lee, “Finding Reliable Recommendations for Trust Model”, The 7th International Conference on Web Information Systems Engineering (WISE 2006). LNCS 4255, ISBN 3-540-48105-2. pp 375-386. (LNCS, SCIE)

[9] Weiwei Yuan, Donghai Guan, Sungyoung Lee, Young-Koo Lee, and Heejo Lee, “Filtering out Unfair Recommendations for Trust Model in Ubiquitous Environments”, Second International Conference on Information Systems

Security (ICISS 2006) 17-21 December 2006, Kolkata, India. LNCS 4332, ISBN 3-540-68962-1, pp 357-360. (LNCS, SCIE)

[10] Donghai Guan, Weiwei Yuan, Andrey Gavrilov, Young-Koo Lee, Sungyoung Lee and Sang Man Han, "Using Fuzzy Decision Tree to Handle Uncertainty in Context Deduction", 2006 International Conference on Intelligent Computing, ISBN: 978-3-540-37274-5, ISSN: 0302-9743, LNAI 4114, pp. 63-72. (LNAI, SCIE)

[11] Donghai Guan, Weiwei Yuan, Young-Koo Lee and Sungyoung Lee, "Utilizing a Hierarchical Method to Deal with Uncertainty in Context-aware Systems", 2006 International Conference on Intelligent Computing, ISBN: 978-3-540-37255-4, ISSN: 0170-8643, LNCIS 344, pp. 741-746. (LNAI, SCIE)

[12] Le Xuan Hung, Hassan Jammeel, Seong Jin Cho, Yuan Weiwei, Sungyoung Lee and Young-Koo Lee, "A Trust Model for Uncertainty in Ubiquitous Environments", IEEE Intelligence and Security Informatics Conference (ISI 2006). LNCS: 3975 pp. 755-757. (LNCS, SCIE)

## Patent

[13] Sungyoung Lee, Young-Koo Lee, Yuan Weiwei, "Behavior Based Method And System for Filtering out unfair rating for Trust Models", US Patent No. 12/494,446, 2009.06.30

## Conferences

[14] Weiwei Yuan, Donghai Guan, and Sungyoung Lee, “Trust Management for Ubiquitous Healthcare,” Parallel and Distributed Processing with Applications, 2008. ISPA '08. International Symposium on, 2008, pp. 63-70.

[15] Weiwei Yuan, Donghai Guan, Sungyoung Lee, and Heejo Lee, “Using Reputation System in Ubiquitous Healthcare”, The 9th IEEE International Conference on e-Health Networking, Application & Services (Healthcom 2007), Taipei, June 19-22, 2007. ISBN: 1-4244-0942-x, pp 182-186.

[16] Weiwei Yuan, Donghai Guan, Sungyoung Lee, and Young-Koo Lee, “The Role of Trust in Ubiquitous Healthcare”, the 9th IEEE International Conference on e-Health Networking, Application & Services (Healthcom 2007), Taipei, June 19-22, 2007. ISBN: 1-4244-0942-x, pp 312-315.

[17] Weiwei Yuan, Donghai Guan, Sungyoung Lee and Young-Koo Lee, “A Reputation System based on Computing with Words”, International Wireless Communications and Mobile Computing Conference 2007 (IWCMC 2007),

August 12-16, 2007, Honolulu, Hawaii. ISBN:978-1-59593-695-0. pp, 132-137

[18] Weiwei Yuan, Donghai Guan, Sungyoung Lee, “The Role of Reputation in Ubiquitous Healthcare System”, the 27th KIPS Spring Conference, Korea. pp. 847-848.

[19] Weiwei Yuan, Donghai Guan, Sungyoung Lee and Heejo Lee, “Bayesian Memory-Based Reputation System”, International Mobile Multimedia Communications Conference (MobiMedia 2007), ISBN:978-963-06-2670-5, Article No. 9.

[20] Weiwei Yuan, Donghai Guan, Sungyoung Lee, and Youngkoo Lee, “A Context-Based Architecture for Reliable Trust Model in Ubiquitous Environments”, The 14th IEEE International Conference on Networks (ICON2006), Singapore, Sep 13-15, 2006. ISBN: 0-7803-9746-0, pp236-240.

[21] Weiwei Yuan, Donghai Guan, Le Xuan Hung, Sungyoung Lee, and Youngkoo Lee, “A Trust Model with Dynamic Decision Making For Ubiquitous Environments”, The 14th IEEE International Conference on Networks (ICON2006), Singapore, Sep 13-15, 2006. ISBN: 0-7803-9746-0, pp230-235.

[22] Donghai Guan, Weiwei Yuan, Young-Koo Lee, and Sungyoung Lee, “Training data selection based on Fuzzy C-means”, Proc. of Fuzz-IEEE (WCCI2008), Hong Kong, China. ISBN: 978-1-4244-1818-3, pp. 761-765.

[23] Donghai Guan, Weiwei Yuan, Young-Koo Lee, and Sungyoung Lee, “Semi-supervised Nearest Neighbor Editing”, Proc. of IJCNN-IEEE (WCCI2008), Hong Kong, China. ISBN: 978-1-4244-1820-6, pp. 1183-1187.

[24] Donghai Guan, Weiwei Yuan, Sungyoung Lee and Young-Koo Lee, “Context Selection and Reasoning in Ubiquitous Computing”, The 2007 International Conference on Intelligent Pervasive Computing (IPC-07), October 11th ~ 13th, 2007, in Jeju Island, Korea, ISBN: 978-0-7695-3006-0, pp. 184-187

[25] Donghai Guan, Weiwei Yuan, Seong Jin Cho, Andrey Gavrilov, Young-Koo Lee, Sungyoung Lee:, "Devising an Information Gain-based Reasoning Engine for Context-aware Ubiquitous Computing Middleware", Proc. of International Conference on Ubiquitous Intelligence and Computing (UIC 2007, LNCS), Hong Kong, China, July, 2007. ISBN: 978-3-540-73548-9, pp. 849-857

[26] Donghai Guan, Andrey V. Gavrilov, Weiwei Yuan, Young-Koo Lee and Sungyoung Lee, "A Novel Hybrid Neural Network for Data Clustering", The 2007 International Conference on Machine Learning, Models, Technologies

and Applications, WorldComp 2007, June 25-28, Las Vegas, USA. ISBN 1-60132-027-2. pp. 284-288

[27] Donghai Guan, Weiwei Yuan, Young-Koo Lee, Andrey Gavrilov and Sungyoung Lee, "Combining Multi-layer Perceptron and K-means for Data Clustering with Background Knowledge", The 2007 International Conference on Intelligent Computing (ICIC2007, Springer), August 21-24, Qingdao, China. ISSN 1865-0929 (Print) 1865-0937 (Online). pp. 1220-1226

[28] Donghai Guan, Weiwei Yuan, Young-Koo Lee, Andrey Gavrilov and Sungyoung Lee, "Data Selection Based on Fuzzy Clustering", The 12th International Conference on Fuzzy Theory & Technology (JCIS 2007), July 18-24, USA. DOI No: 10.1142/9789812709677_0174, Source: INFORMATION SCIENCES 2007, pp 1231-1237.

[29] Donghai Guan, Weiwei Yuan, Young-Koo Lee, Andrey Gavrilov and Sungyoung Lee, "Activity Recognition Based on Semi-supervised Learning", The 13th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications, August 21-24, 2007, Korea. ISSN: 1533-2306.ISBN: 978-0-7695-2975-2.pp. 469-475

[30] Donghai Guan, Andrey Gavrilov, Weiwei Yuan, Sungyoung Lee and Young-Koo Lee, "Data Clustering Using Hybrid Neural Network", the 27th KIPS Spring Conference, Korea. pp. 457-458.

[31] Brian J. d'Auriol, Pramod Chikkappaiah, Weiwei Yuan, Sungyoung Lee and Young-Koo Lee, "Query Responsive Awareness Software: Inventory Control Case Study", ACM 2nd international conference on Ubiquitous information management and communication (ICUIMC 2008), ISBN:978-1-59593-993-7, pp. 520-524

[32] Hassan Jameel, Riaz Ahmed Shaikh, Le Xuan Hung, Yuan WeiWei, S. M. K. Raazi, N. T. Canh, S. Lee, H. Lee, Y. Son, and Mi. Fernandes, "Image-Feature based Human Identification Protocols on Limited Display Devices", 9th International Workshop on Information Security Applications (WISA 2008), Jeju Island, Korea, Sep 2008 .

[33] L.X. Hung, R.A. Shaikh, H. Jameel, S.M.K.R. Raazi, Y. Weiwei, N.T. Canh, P.T.H. Truc, S. Lee, Y. Son, and M. Fernandes, "Activity-Oriented Access Control for Ubiquitous Environments," Consumer Communications and Networking Conference, 2009. CCNC 2009. 6th IEEE, 2009, pp. 1-5.