The influence of trust on sharing information

Danielle Hyunsook Lee
hyl12@pitt.edu

ABSTRACT

In collaborative filtering recommender systems, there is little room for users to get involved in the choice of their peer group. It leaves users defenseless against various spamming or "shilling" attacks. Other social Web-based systems, however, allow users to self-select trustworthy peers and build a network of trust. We argue that users self-defined networks of trust could be valuable to increase the quality of recommendation in CF systems. To prove the feasibility of this idea we examined how similar are interests of users connected by a self-defined relationship of trust in two collaborative tagging systems *Delicious* and *Citeulike*. Interest similarity was measured by similarity of items and sources they share and tags they use. Our study shows that users connected by a network of trust exhibit significantly higher similarity on all expored levels (items, sources, and tags) than non-connected users. This similarity is highest for directly connected users and decreases with the increase of distance between users. The data also hints that connected users tend to share specific rather than general interests, as shown by rare information items and tags they have in common.

1. INTRODUCTION

Recommender systems powered by collaborative filtering (CF) technologies become a feature of our life. Such popular systems as Amazon.com, Netflix, Last.fm, and Google News use Collaborative filtering to recommend us products to buy, movies to watch, music to listen and news to read. The power of this technology is based on a relatively simple idea: starting with a target user's rating, find a cohort of users who have similar interests and recommend items favored by this cohort to the target user. While generations of collaborative recommender systems have proven the effectiveness of this approach, modern recommender systems share one important flaw: they provided no ways for the user to control the recommendation mechanism. The users of CF systems are not involved in the process of choosing the peer group and generating recommendation. Unlike early-day "push" and "pull" collaborative filtering systems [13, 18] which connected people openly, modern collaborative filtering systems do not even expose information about the people whose ranking was used to generate the recommendations. A user can't ascertain whether the users in the peer cohort selected by the recommender system really do have similar interests and whether they are simply trustworthy. As a result, recommendation in modern collaborative filtering system is provided by users who are neither known, nor trusted by the end users.

The lack of control over the selection of peers has disadvantages from both sides. Target users can neither add specific known and trustworthy persons (someone they consult in real life) to the set of their peers, nor can they exclude some strange or suspicious "peers" from the set picked up by the system. The latter inability leaves users defenseless against various spamming or "shilling" attacks in collaborative recommenders. Thus, while growing more powerful over the years, modern CF systems have lost what they started with: *trust*.

At the same time, the issue of personal relations and trust has become the driving force of a new generation of social systems brought to life by Web 2.0. A range of Web 2.0 systems such as *LlinkedIn*, *Flickr*, *Delicious*, *Citeulike* etc., provide various kind of social linking, enabling their user to pick known and trusted users and add them to their list of connections. These self-defined links between users establish a rich network of trust, which is, in turn, used to propagate various kinds of information. Explicitly or implicitly the users in these trusted networks recommend each other group to join, events to attend, bookmarks to explore, or research papers to read.

The success of social linking and bookmaking systems that allow users to build their networks of trust, stresses a fact forgotten by modern CF systems: the source of the recommendation is an important criterion for judging the quality of recommendations [3]. Given that, it is natural to expect some kind of merger between social linking a collaborative filtering technology: a new generation of *trust-based recommender systems*, which will use self-defined social networks of trust to improve the quality of CF systems and the satisfaction of their users. Some pioneer works in this direction already appeared [11, 12, 13, 17, 24]

To prove that trust-based recommenders are more than a speculation, some important assumptions have to be checked. Is it true that connected users in the networks of trust share not only trust, but also some common interests? Is it true that information can flow along these networks, i.e., the choices made by users are affected by the choices of users they trust? If these assumptions appear to be invalid, then such networks of trust will be useless for CF mechanisms.

The goal of this paper is to test these assumptions. Using real life data collected from two social Web systems, *Delicious* and *Citeulike*, we examined several important properties of self-defined trust networks. We investigated how similar are users interests in these networks, the extent to which the amount of similar information collected by users depends of the strength of their connection, and ultimately, how feasible it may be to use a network of trust as a source of personalized recommendation.

The term 'trust' as used in this paper may not be an exact match with the general usage of 'trust' as defined in the sociology. The social relationship used in this paper is defined unilaterally, simply indicating user trust in the usefulness of information provided by connected individual. It is not trust through personal interaction or emotional support (for instance, connected with an expectation of obligation, morality or responsibility [14]). However, no better or more precise term for this relationship existed; since referred users are deemed "trustworthy" by the target user in terms of information collection, the term 'trust' was selected for use herein. Furthermore, the term 'trust' as defined in the Webster's Third New International Dictionary meets our interpretation of 'trust.' Its definitions for the term are "a confident dependence on the character, ability, strength, or truth of someone or something," "confident anticipation," and "a charge or duty imposed in faith and confidence or as a condition of some relationship" [14].

This paper is organized as follows. In Section 2, we review the literature surrounding CF recommendation, social networks existing on the Web, and social tagging. The way to collect data, the description of data and the kinds of relationship are explained in Section 3. Section 4 concerns the variables and the hypotheses we tested and Section 5 presents the results. We conclude with a discussion of this study and suggestions for future research.

2. RELATED WORK

2.1 CF Technology and its Problems

The collaborative filtering technology emerged as an attempt to automate the word-of-mouth in the age of Internet. The technology proved its worth in recommending taste-based items such as movies, jokes, music, etc. where the preference is hard to be appreciated by the content. It became popular for its ability to recommend serendipitous and diverse information. The popularity, however, revealed some problems associated with CF approach. CF appeared to be not well-protected against malicious users who try to harm the system or to make a profit by gamming the system. For example, by copying the whole user profile, a malicious user is perceived by the system to be a perfect peer user and the products added by him are therefore recommended to the target user [9, 12, 15]. Even without malicious users the quality of recommendation can be affected by peculiar users with unusual interests [18]. Moreover, since recommenders have to compare all other users in order to find the peer group, the computation requires substantial off-line process [11]. Finaly, users who do not have sufficient ratings are not able to receive reliable recommendations [18]. These CF-related problems occur in part because the recommender systems make a choice of peer group purely by similarity computation, with no way for the target users to affect the recommendation process.

Several research teams attempted to exploit trust between users to resolve some of the cited problems of CF technology. Massa and Avesani's study [11] showed that a user's trust network can solve the ad-hoc user problem, improve recommendation prediction and attenuate the computational complexity. Using *Epinions* data set, a trust-based technology generated more precise recommendations than CF technology. In addition, for users with 4 ratings, trust-based technology could make recommendation for 66% of the users, while CF could only make recommendations for 14% of the users with a higher margin of error. As a result, they suggested that collecting a few trust statements is more effective, in term of coverage and error, than collecting an equivalent number of ratings . Another study indicated that a trusted network decreases the recommendation error and increases the accuracy as well [17]. However these studies rely on inferred trust, not users-defined trust. For users with a unique taste, their own trusted network could increase the satisfaction of recommendation, since they are able to know where the information cames from [22]. The recommendations made by friends were known to be frequently better and more useful than the recommendation made by systems [21]. In this context, our study presented below expands the current body of work by exploring the value of self-defined networks of trust as a source of information for recommendation.

2.2 Social Relationship on the Web

In early CF-related research, social networking was an important part of recommendation process. The first collaborative filtering system, *Tapestry* was based on explicit social connections and allowed users to retrieve personalized contents, using annotations added by their friends or colleagues [18]. Another pioneering CF project [13] combined explicit social connections with active "push" approach: users could directly send interesting research papers to other colleagues. However, these pioneer systems relied on exchange of information within a "small world" and found it difficult to retain users and to keep them actively contributing to the recommendation process. As the CF-related algorithms became mature, automatic recommendations by computations became dominant.

More recently, the problems of CF technology caused some researchers to re-asses the value of explicit social connections as a source of information for reliable recommendation. To prove the feasibility of this approach, several research teams started with checking the main assumption: do users linked by self-defined networks of trust have similar interests.

Ziegler and Golbeck [24] compared interest similarity between people in a trusted network in traditional CF context. They used information regarding users and the user's trust ratings in the book and movie recommendation. For the book data set, rather than using each information item, they grouped the items by topics, using an existing taxonomy. Then, they built topic-based user profiles, to lessen the data sparsity problem. The closeness of the user profiles in the trusted network was assessed. However, because this study focused on the similarity of users' *topics* of interests derived from the taxonomy, not from the *items* per se, it was hard to see a clear picture of information similarity on the item level. As the authors pointed, the similarity is highly dependent on the taxonomy's design. Yet, in the second study with movie data set, they compared user similarity on the level of individual items and found that the average ratings of each movie became more similar as the trust values between two users increased [24].

Singla and Richardson (2008) tested the relationship between instant messenger logs and the similarity of search queries. They were able to demonstrate that search interests of people who exchanged instant messages frequently were more similar than interests of random pairs. Moreover, the longer they talked, the more similar they were [20].

Another trust-related research [1] explored terms on the Stanford and MIT personal home pages and users' social connections. The terms in homepages, in-links, out-links and mailing lists were analyzed to see how the information similarity predicts the friendship connections. All these four kinds of information appeared to be similar for socially connected users. There was also proportional relationship between information dissimilarity and the relationship distance [1].

Our work presented below was motivated the same goal: to assess interest similarity between users connected by relationship of trust. Expanding the research cited above, we attempted to explore this topic in a very different context: collaborative tagging systems.

2.3 Cognitive Aspects of Tag

Tags are descriptive words selected by users to describe items in collaborative information sharing systems. Tags were found to be helpful for organizing, retrieving and sharing information. The collection of the tags is often called 'folksonomy' (which a compound word of 'folk' and 'taxonomy' [7]) equaling it to a collection of conceptual and structured knowledge created by people. According to the analysis of tags sharing in Connotea, around 23% of tags out of 3359 unique tags were shared between users [10]. Although it was preliminary result based on initial usage of the system, it shows that tags are being utilized not only as a way to organize information for personal goods, but as a good way to help discover useful information shared by others.

Two related papers explored the motivation of tagging behavior in terms of social presence [2, 16]. They defined target audiences of users' tags into three groups, self, family & friends and generic public. According to the kind of target audience in users' mind, they compared the correlation of social presence and the number of tags in Flickr. The tagging motivations for themselves and generic public significantly correlated with the number of tags, but there was no correlation between the number of tags and the tagging motivation by friends and family. That is to say, users don't tag for the sake of their friends and family [16], but for the general audiences. However, as pointed by the authors, the data cannot explain the general motivation of tagging behavior. The phenomenon could be caused by the nature of pictures. Pictures are main media to communicate with friends and family in person. Therefore, there could be another way to share the information of their pictures such as phone call, email or conversation in offline.

User tagging behavior can serve as rich evidence about user interests and in this capacity was applied to construct user profiles for personalized information access. Compared with item ratings, item tags not only indicate user interest, but also show from which aspects the tagged item is interesting for the user [8; 23]. Another study [8] using *Delicious* data set explored the correlation between users' tags and content tags. The authors also explored how well personal tag sets and social contacts' tag sets represent their preferences. As expected, personal tag were more effective to express user preferences. However, user profiles based solely on tag sets were not good enough to generate recommendation (the best precision was 21%).

In the domain of music recommendation, tag-based user profiles were also found not good enough to produce recommendations. In Last.fm, precision of recommendation with tag-based user profiles was significantly lower than typical CF-based recommendation. However, once tags were correlated semantically with query terms, the precision of recommendation with tag-based profiles was significantly improved [4].

3. DATA COLLECTION

3.1 Data Sets

As a source of data for our study we selected two different collaborative tagging systems: *Delicious* and *Citeulike*. *Delicious* is a generic social bookmarking system while *Citeulike* is a narrow-focused system for sharing bibliographic references. To pick up initial set of *Delicious* users, we visited this site randomly on the first and second week of November 2007. 1150 users who posted new bookmarks at that time were chosen. For each user we collected the bookmarks, the tags, and user *networks* (connected users). Similarly, we randomly visited *Citeulike* in September and October of 2008. 1365 users of *Citeulike* who posted new articles at the time of visit were picked. The information collected for each user included the bibliography (article title, list of authors, journal name, publication year, etc.), the tags and the *watchlists* (connected users).

Table 1. Data Summary of *Delicious*

Total no. of users	4810
Total no. of distinct items (bookmarks)	109103
Average no. of items per user	26.27
Total no. of distinct tags	43511
Average no. of tags per user	93.42
Total no. of unidirectional relation	5012
Total no. of reciprocal relation	532

After collecting a group of initial users, we collected data of their trusted connections. In collaborative tagging systems explicit connections between users are of special nature. In some sense, they bear more "trust" than the connections between friends in social networking systems. In *Delicious*, if a user considers bookmarks of another user interesting and trustworthy, she can add this user to her 'network.' As a result, the whole list of the connected user's bookmarks is automatically shown to her. Likewise, in *Citeulike*, users can directly connect to other users who have interesting bibliography by adding them 'watch list.' Then the system displays the whole bibliography of watched users.

To diversify data collection, especially, in *Citeulike*, we used several kinds of links to reach and crawled other users. All of the other users who tagged the same articles which already collected users were added to their list were chosen; accordingly, their information and the tags were crawled. Trusted connections of these users were crawled as well (i.e., we crawled two steps away from the original sets).

Table 2. Data Summary of *Citeulike*

Total no. of users	21077
Total no. of distinct items (Papers)	449824
Average no. of items per user	28.69
Total no. of distinct tags	136274
Average no. of tags per user	12.91
Total no. of unidirectional relation	11295
Total no. of reciprocal relation	93

Table 1 and Table 2 show the descriptive statistics of each data set. As the data implies, out of total number of distinct items, the percentage of items which are stored by more than one user is larger in *Citeulike* data set (17.22%) than *Delicious* (7.91%).

3.2 The Networks of Trust

In this paper, we interpret user's act of connecting to other users (by adding this user to the network of watch list) as a sign that she likes the focus and trust the quality of the added user's bookmarks or references and wants to have direct access to them continuously in future. Thus, networking in *Delicious* and watching in *Citeulike* could be considered as evidence that connected users are trustworthy to the original user in terms of information collection.

We distinguish two kinds of trusted connections – unidirectional and reciprocal. In both *Delicious* and *Citeulike*, the act of adding another user to the 'network' or 'watch list' is unidirectional (which is different from social networking systems). If user A added user B to her network, it does not imply that user A will be added to B's network at the same time. The users in A's network decide independently whether to add A to their networks. For example, user B may not have A in his network. In this case, we call the relationship between A and B as 'unidirectional'. Another user C in A's network may add A to his network as well (Figure 1). We call this relationship as 'reciprocal'.

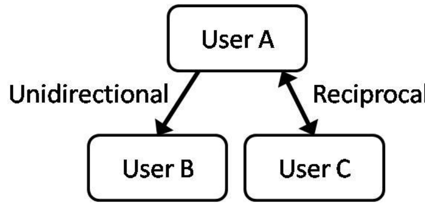

Figure 1. Directions of relation in the center of user A

We also distinguish direct and indirect connection and use them it to investigate the transitivity of common interests in the networks of trust. In our study we explored three distances between users in trust networks: *direct*, *one hop* and *two hops*. In the above example, user A and user B are in 'direct' relationship. If user B is trusting user D, user A and user D are in 'one hop distance' unidirectional relationship (Figure 2). If user E belongs to the network (or watchlist) of user D, users A and E are in 'two hop distance' unidirectional relationship. This distance can be applied to the reciprocal relationship as well.

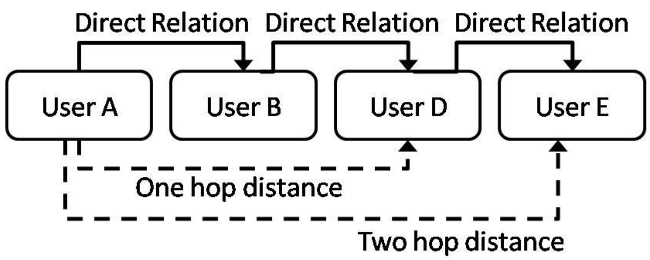

Figure 2. Relation distance in the center of user A

As Table 1 and Table 2 show, our data set included a good number of unidirectional connections, although the average number of connections per user is much smaller than in social networking systems: just about one per user in *Delicious* and a half per user in *Citeulike*. The average number, however, does not provide a good picture since connections are distributed unevenly. (see Figure 3 for the distribution of unidirectional relationships on the center of trusting users*;* Figure 4 for the distribution of reciprocal relationship on the center of trusting users. The reciprocal relationships were counted twice from both directions).

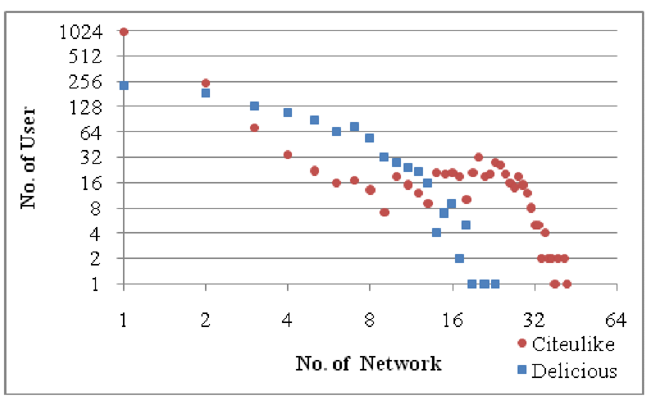

Figure 3. Distribution of Unidirectional Relation

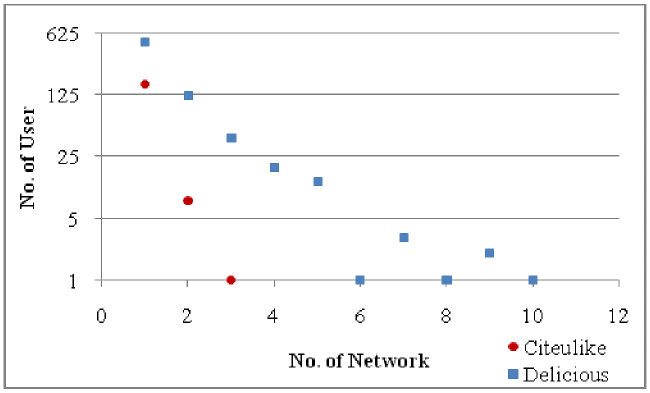

Figure 4. Distribution of Reciprocal Relation

4. DATA ANALYSIS

The similarity of user interests in collaborative tagging systems can be assessed in two ways: the similarity of shared information and the similarity of used tags. In this study we explored both aspects of interest similarity in the networks of trust: how similar is the information shared by people in trusted network and how similar are their tags.

We assessed item-based and tag-based similarity by counting the number of shared information *items*, information *sources* and *tags*. Information items are academic papers in *Citeulike* and bookmarks in *Delicious*. Information sources represent *origins* of information items. We assume that information items coming from the same source could be similar and sometimes even duplicating. We hypothesized that the users who share the same interests may not necessarily agree about specific shared items, but demonstrate higher agreement on the level of item sources. For example, in *Citeulike*, paper authors and journals (or conferences) can be considered as sources for references. In our *Citeulike* dataset, we considered authorship only since it is more reliable and easy to track. In *Delicious* we considered root address of each bookmark (i.e., Web site it came from) as its source. Finally, in tag similarity, we considered two aspects – *micro* level and *macro* level tag similarity. Micro level represents tag similarity for a common article and macro level represents the similarity of the overall collections of tags of two users.

4.1 Dependent Variables

Since sizes of item and tag collections varied dramatically from user to user, we had to examine both absolute and relative similarity measures. That is to say, in order to measure between-users' information similarity, we not only used absolute numbers (i.e., number of common bookmarks or tags), but we also compared *relative* (normalized) numbers: proportion of shared items in respect to the collections of connected users. We used three meaningful relative similarity measures as dependent variables: how much the information of a trusting user is influenced by the trusted user (in-link power), how much the information of a trusted user affects the trusting user (out-link power) and how which fraction of joint information is overlapped overall (overall power). Figure 5 and the following equations explain the meaning of these measures.

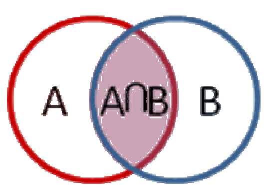

Figure 5. Information Overlap

$$\text{Inlink Power} = (A \cap B)/A \qquad \text{eq. (1)}$$

$$\text{Outlink Power} = (A \cap B)/B \qquad \text{eq. (2)}$$

$$\text{Overall Power} = (A \cap B)/(A \cup B) \qquad \text{eq. (3)}$$

If user A added user B to her trusted network (i.e, A points to B), the inlink power (impact) of the user B for the user A is the fraction of the A's shared information with user B in the user A's collection. The outlink power of User B is the fraction of shared information with user A in respect to user B's collection. The overall power is the fraction of shared information in the joint information space of both users.

4.2 Hypotheses

4.2.1 Information Similarity in Trust Network

For the information similarity in trusted network, four hypotheses were assessed:

H1. Users connected by direct or indirect relationships of trust have more similar information (information item and the sources) than a non-connected pairs.
H2. Users in reciprocal relations have more similar information (information item and the sources) than users in unidirectional relations.
H3. Users with more connections in their trusted network are able to assemble more resources.
H4. The more popular items are, the more they are shared by connected users (in contrast to a non-connected pair).

4.2.2 Tag Similarity in Trust Network

For tag similarity of trusted network two hypotheses were tested.

H5. Users connected by a relation of trust have more similar tags on macro- and micro-level than non-connected users.
H6. Users who are connected as trusted network share more relatively less popular tags than popular tags.

5. THE RESULTS

5.1 Information Patterns

Before examining tests for the hypotheses regarding information similarity, we examined the information sharing patterns. As we found out, users in our datasets had very different interests and the overall level of information sharing was quite low. Out of 11,295 direct unidirectional pairs in *Citeulike*, 9,087 pairs did not have any common reference, 10,333 pairs did not use the same tag on the same resource (micro level tag similarity), and 2,798 have no common tags at all (macro level tag similarity). The remaining users shared one or more items or tags. Figure 6 and Figure 7 show how the amounts of shared information is distributed over the users for direct unidirectional and direct reciprocal relations (after excluding the pairs with zero overlap). As some other distributions in collaborative tagging systems, it exhibits features of exponential distribution (appearing as straight line in log-log coordinates).

Similarly, out of 5,012 direct unidirectional relations in *Delicious*, 4558 pairs did not share any bookmark, 4697 pairs did not assign the common tags for the same items (micro-level) and 899 pairs had no common tags in their tag collections (macro-level). Figure 8 and Figure 9 represent the distribution of sharing patterns in direct unidirectional relation and direct reciprocal relations, respectively.

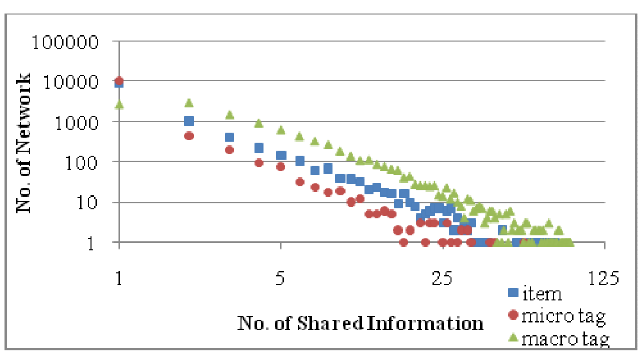

Figure 6. Distribution of Shared Information
(Direct Unidirectional Relationships, *Citeulike*)

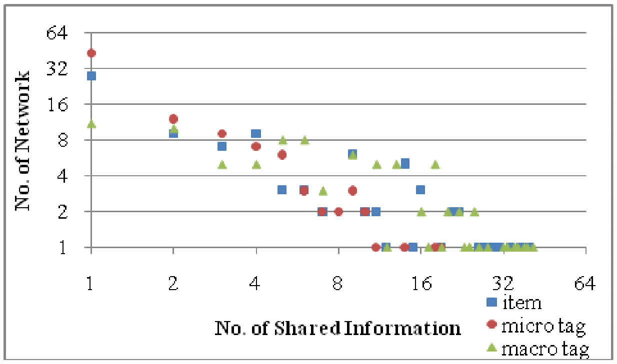

Figure 7. Distribution of Shared Information (Direct Reciprocal Relationships, *Citeulike*)

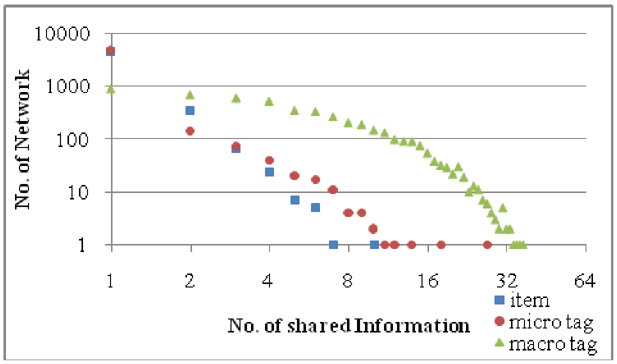

Figure 8. Distribution of Shared Information (Direct Unidirectional Relationships, *Delicious*)

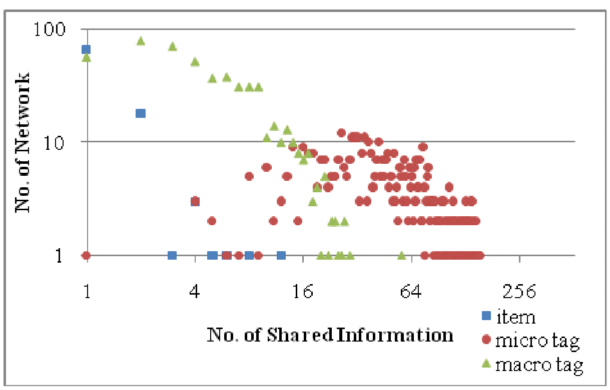

Figure 9. Distribution of Shared Information (Direct Reciprocal Relationships, *Delicious*)

5.2 Information Similarity in Trust Networks

5.2.1 Information sharing pattern in trusted network

To test whether users connected by direct or distant links of trust share more information than non-connected pairs (H1), we compared both absolute numbers of shared information items and their normalized numbers (inlink, outlink, and overall powers).

Table 3. The Number of the Common Information (Unidirectional Relation)

		Direct	1hop	2hops	No Rel.
Citeulike	Items	**.81**	**.22**	**.15**	**.00**
		F (3, 154667) = 2133.13, p < .001, η^2 = .040			
	Sources	**65.57**	**38.72**	**34.70**	**.07**

		F (3, 154667) = 109.814, p < .001, η^2 = .038			
Delicious	Items	**.12**	**.03**	**.02**	**.01**
		F (3, 84147) = 836.00, p < .001, η^2 = .029			
	Sources	**.52**	**.31**	**.25**	**.13**
		F (3, 84147) = 1036.927, p < .001, η^2 = .046			

First, we explored the number of shared items and sources. Table 3 shows mean numbers of shared items and sources for direct and distant unidirectional relationship on contrast to a non-related pair of users (which we can interpret as infinite distance). As we can see, at average, direct pairs share the largest number of items and sources in both *Citeulike* and *Delicious.* The numbers are decreasing with the increase of distance in the network of trust achieving its minimum for non-connected pairs (which can be considered as infinite distance). This difference appeared to be significant for all four rows of the table (i.e., for both items and sources and for both systems). This is the evidence that users connected in a network of trust do have significantly more similar interests than non-connected users. We can also consider it as an evidence of information propagation along a network of trust, although impressive similarity on the source level (which are hard to propagate!) hints that interest similarity may play a more important role than propagation in the observed phenomenon.

Table 4. The Number of the Common Information (Reciprocal Relation)

		Direct	1hop	2hop
Citeulike	Items	**8.35**	**1.77**	**.73**
		F (2, 1457) = 137.71, p < .001, η^2 = .159		
	Sources	**93.02**	**62.32**	**32.56**
		F (2, 1457) = 9.483, p < .001, η^2 = .013		
Delicious	Items	**.28**	**.04**	**.02**
		F (2, 17132) = 419.148, p < .001, η^2 =.046		
	Sources	**.79**	**.39**	**.27**
		F (2, 17132) = 312.296, p < .001, η^2 = .035		

As First, we explored the number of shared items and sources. Table 3 shows mean numbers of shared items and sources for direct and distant unidirectional relationship on contrast to a non-related pair of users (which we can interpret as infinite distance). As we can see, at average, direct pairs share the largest number of items and sources in both *Citeulike* and *Delicious.* The numbers are decreasing with the increase of distance in the network of trust achieving its minimum for non-connected pairs (which can be considered as infinite distance). This difference appeared to be significant for all four rows of the table (i.e., for both items and sources and for both systems). This is the evidence that users connected in a network of trust do have significantly more similar interests than non-connected users. We can also consider it as an evidence of information propagation along a network of trust, although impressive similarity on the source level (which are hard to propagate!) hints that interest similarity may play a more important role than propagation in the observed phenomenon.

Table 4 show, reciprocal relationships, in both *Citeulike* and *Delicious*, exhibit the same pattern, also with significant differences between columns in the number of shared information items and sources.

Second, we explored differences between relative similarity measures – fractions of shared items and sources for unidirectional relationships in *Citeulike* (Table 5), unidirectional relationships in *Delicious* (In addition to demonstrating a clear connection between item and source-level similarity and user closeness in a network of trust, the data shown above allows to make three interesting observations. First, as we expected, between-user similarity on the level of *sources* is much larger than similarity on the level of *items* for both systems. For example, in *Citeulike,* the inlink power similarity of items in direct relation is 2.01% while inlink power similarity of sources in the same direct relation was 6.37%. Second, both absolute and relative similarities are much larger in *Citeulike*, a very specialized sharing system, than in *Delicious*, a very general sharing system. Third, both absolute and relative similarities are pair-wise larger for reciprocal than for unidirectional connections for all distance levels and both systems. This difference is most pronounced in relative form reaching its highest level for direct reciprocal relations in *Citeulike* (6.79% for items and 13.01% for sources). Next section examines the difference between reciprocal and unidirectional connections in details and checks its significance.

Table 6), reciprocal relationships in *Citeulike* (Table 7), and reciprocal relationships in *Delicious* (Table 8). In all four cases, the differences for all relative measures (inlink power, outlink power, and overall power) – were significant among the network distances. Same pattern can be observed in the case of relative similarity measures: directly related users have the largest fraction of shared items and sources, this fraction decreases with the increase of the distance between users and reaches its minimal level for not connected users (infinite distance).

Table 5. Similarity Powers of information items in *Citeulike* (Unidirectional Relations)

		Direct	1hop	2hop	No Rel.
Items	Inlink	**2.01%**	**0.16%**	**0.11%**	**0.01%**
		F (3, 154667) = 2214.714, $p < .001$, $\eta^2 = .042$			
	Outlink	**0.83%**	**0.60%**	**0.44%**	**0.00%**
		F (3, 154667) = 686.662, $p < .001$, $\eta^2 = .013$			
	Overall	**0.34%**	**0.07%**	**0.05%**	**0.00%**
		F (3, 154667) = 2590.056, $p < .001$, $\eta^2 = .048$			
Sources	Inlink	**6.37%**	**3.49%**	**3.32%**	**0.05%**
		F (3, 154667) = 174.132, $p < .001$, $\eta^2 = .062$			
	Outlink	**7.51%**	**2.91%**	**2.57%**	**0.07%**
		F (3, 154667) = 273.216, $p < .001$, $\eta^2 = .089$			
	Overall	**2.59%**	**1.26%**	**1.26%**	**0.01%**
		F (3, 154667) = 166.942, $p < .001$, $\eta^2 = .056$			

In addition to demonstrating a clear connection between item and source-level similarity and user closeness in a network of trust, the data shown above allows to make three interesting observations. First, as we expected, between-user similarity on the level of *sources* is much larger than similarity on the level of *items* for both systems. For example, in *Citeulike,* the inlink power similarity of items in direct relation is 2.01% while inlink power similarity of sources in the same direct relation was 6.37%. Second, both absolute and relative similarities are much larger in *Citeulike*, a very specialized sharing system, than in *Delicious*, a very general sharing system. Third, both absolute and relative similarities are pair-wise larger for reciprocal than for unidirectional connections for all distance levels and both systems. This difference is most pronounced in relative form reaching its highest level for direct reciprocal relations in *Citeulike* (6.79% for items and 13.01% for sources). Next section examines the difference between reciprocal and unidirectional connections in details and checks its significance.

Table 6. Similarity Powers of information items in *Delicious* (Unidirectional Relations)

		Direct	1hop	2hop	No Rel.
Items	Inlink	**0.50%**	**0.11%**	**0.07%**	**0.02%**
		F (3, 84147) = 677.624, $p < .001$, $\eta^2 =.024$			
	Outlink	**0.57%**	**0.12%**	**0.07%**	**0.02%**
		F (3, 84147) = 630.345, $p < .001$, $\eta^2 =.022$			
	Overall	**0.24%**	**0.05%**	**0.03%**	**0.01%**
		F (3, 84147) = 838.357, $p < .001$, $\eta^2 =.029$			
Sources	Inlink	**2.45%**	**1.70%**	**1.30%**	**0.53%**
		F (3, 84147) = 887.751, $p < .001$, $\eta^2 =.040$			
	Outlink	**2.81%**	**1.59%**	**1.23%**	**0.49%**
		F (3, 84147) = 1150.989, $p < .001$, $\eta^2 =.051$			
	Overall	**1.16%**	**0.72%**	**0.55%**	**0.24%**
		F (3, 84147) = 1179.215, $p < .001$, $\eta^2 =.052$			

Table 7. Similarity Powers of Shared Information in *Citeulike* (Reciprocal Relations)

		Direct	1hop	2hop
Items	Inlink & Outlink	**6.79%**	**1.18%**	**0.38%**
		F (2, 1457) = 155.134, $p < .001$, $\eta^2 =.176$		

	Overall	**2.45%**	**0.35%**	**0.10%**
		F (2, 1457) = 267.477, $p < .001$, η^2 =.269		
Sources	Inlink & Outlink	**13.01%**	**6.25%**	**3.17%**
		F (2, 1457) = 37.966, $p < .001$, η^2 = .050		
	Overall	**4.79%**	**2.29%**	**1.22%**
		F (2, 1456) = 25.832, $p < .001$, η^2 =.034		

Table 8. Similarity Powers of Shared Information in *Delicious* (Reciprocal Relations)

		Direct	1hop	2hop
Items	Inlink & Outlink	**1.25%**	**0.18%**	**0.08%**
		F (2, 8655) = 337.088, $p < .001$, η^2 = .037		
	Overall	**0.51%**	**0.07%**	**0.03%**
		F (2, 8655) = 426.762, $p < .001$, η^2 =.047		
Sources	Inlink & Outlink	**3.88%**	**2.07%**	**1.53%**
		F (2, 8655) = 150.736, $p < .001$, η^2 =.017		
	Overall	**1.66%**	**0.90%**	**0.65%**
		F (2, 8655) = 196.252, $p < .001$, η^2 =.022		

5.2.2 Unidirectional vs. Reciprocal Relations

To compare the differences of information sharing pattern between unidirectional and reciprocal relations, we started with comparing the number of shared information items and sources, doing it now separately for several distances of relations.

In all three distances in *Citeulike*, the numbers of shared information items in reciprocal relations were significantly larger than in unidirectional relations. For *Delicious* dataset, we found the same pattern (refer to Table 3, First, we explored the number of shared items and sources. Table 3 shows mean numbers of shared items and sources for direct and distant unidirectional relationship on contrast to a non-related pair of users (which we can interpret as infinite distance). As we can see, at average, direct pairs share the largest number of items and sources in both *Citeulike* and *Delicious.* The numbers are decreasing with the increase of distance in the network of trust achieving its minimum for non-connected pairs (which can be considered as infinite distance). This difference appeared to be significant for all four rows of the table (i.e., for both items and sources and for both systems). This is the evidence that users connected in a network of trust do have significantly more similar interests than non-connected users. We can also consider it as an evidence of information propagation along a network of trust, although impressive similarity on the source level (which are hard to propagate!) hints that interest similarity may play a more important role than propagation in the observed phenomenon.

Table 4 and Table 9). Figure 10 displays the differences between unidirectional and reciprocal relations visually.

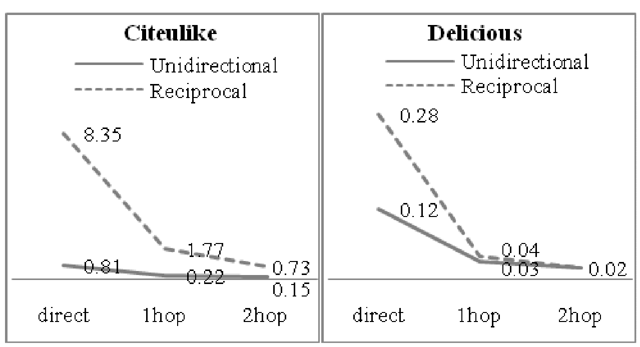

Figure 10. No. of Shared Information Items

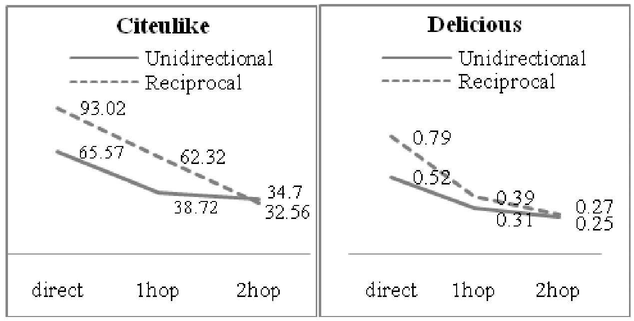

Figure 11. No. of Shared Sources

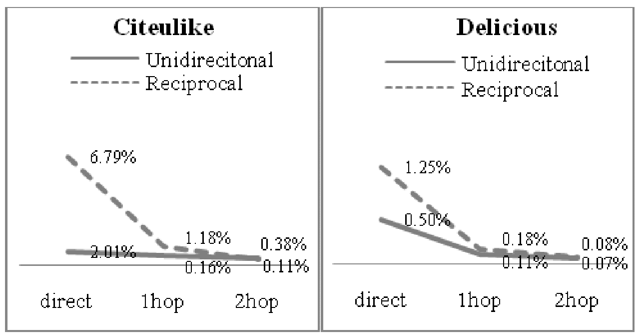

Figure 12. Inlink Powers of Information Items

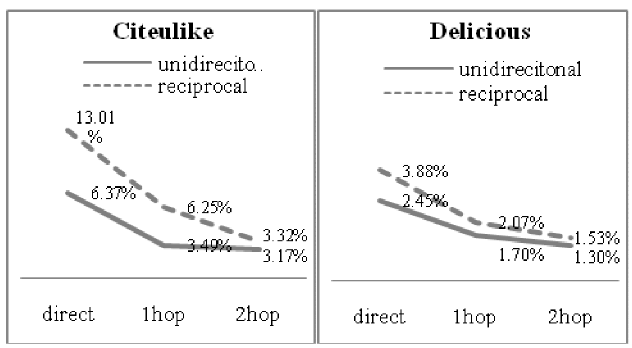

Figure 13. Inlink Powers of Sources

The difference between unidirectional and reciprocal relations in the number of shared sources was also significant in *Delicious* for all three distances (Table 9). For *Citeulike*, however, only difference in 1 hop relation appeared to be significant. Figure 11 shows the mean numbers of shared sources for two kinds of relationship.

Table 9. Test Results of No. of Shared Information Items

	Citeulike			*Delicious*		
	df	*t*-value	Sig.	*df*	*t*-value	Sig.
Direct	11480	-28.22*	p < .001	6075	-8.04*	p < .001
1Hop	20208	-26.10*	p < .001	10938	-4.15*	p < .001
2Hop	32805	-17.73 *	p < .001	24231	-2.10*	p = .036

Table 10. Test Results of No. of Shared Source

	Citeulike	*Delicious*

	df	*t*-value	Sig.	*df*	*t*-value	Sig.
Direct	11480	-1.59*	p =.112	6075	-8.28*	p < .001
1Hop	20208	-2.36*	p =.019	10938	-6.04*	p < .001
2Hop	32805	.288	P =.773	24231	-2.31*	p = .021

Secondly, we checked the significance of observed differences in relative information item similarity (inlink, outlink and overall powers) between reciprocal and unidirectional relations (Figure 12; Table 11 and Table 12). For direct and 1-hop relationship, the differences appeared to be significant for both systems, i.e., users connected by a direct or 1-hop reciprocal relation shared significantly larger fractions of information items than users connected by unidirectional relation. For 2-hops relations the observed difference appeared to be non-significant for four out of six relative similarity measures.

Table 11. Results for Powers of information Items (*Citeulike*)

		Df	*t*-value	Sig.
Inlink Power	Direct	11480	-7.39*	p < .001
	1Hop	20208	-17.12*	p < .001
	2Hop	32805	-8.21*	p < .001
Outlink Power	Direct	11480	-21.51*	p < .001
	1Hop	20208	-2.90*	p = .004
	2Hop	32805	.47	p = .637
Overall Power	Direct	11480	-21.27*	p < .001
	1Hop	20208	-14.67*	p < .001
	2Hop	32805	-5.63*	p < .001

Table 12. Test Results about Powers of information Items (*Delicious*)

		Df	*t*-value	Sig.
Inlink Power	Direct	6075	-8.61*	p < .001
	1Hop	10938	-3.99*	p = .001
	2Hop	24231	-1.06	p = .291
Outlink Power	Direct	6075	-6.54*	p < .001
	1Hop	10938	-3.21*	p < .001
	2Hop	24231	-1.16	p = .245
Overall Power	Direct	6075	-7.80*	p < .001
	1Hop	10938	-3.79*	p < .001
	2Hop	24231	-1.57	p = .116

On the final step we compared relative information source similarity (inlink, outlink and overall powers) for reciprocal and unidirectional relations (Figure 13; Table 13 and Table 14). For *Delicious*, these differences appeared to be significant for all three distances (Table 8 and Table 14). For *Citeulike* relative source similarity was significantly higher for users connected by direct and 1-hop reciprocal relation than for users connected by unidirectional relations of the same distance. No significance difference was observed for users connected by 2-hops relation (Table 7 and Table 13).

Table 13. Test Results about Power of Sources (*Citeulike*)

		df	*t*-value	Sig.
Inlink Power	Direct	11480	-2.41*	p = .016
	1Hop	20208	-4.34*	p < .001
	2Hop	32805	-.73	p = .466
Outlink Power	Direct	11480	-4.54*	p < .001
	1Hop	20208	-2.68*	p = .007
	2Hop	32805	.16	p = .875
Overall Power	Direct	11480	-2.96*	p = .003
	1Hop	20208	-2.73*	p = .006
	2Hop	32805	-.292	p = .770

Table 14. Test Results about Power of Sources (*Delicious*)

		df	*t*-value	Sig.

Inlink Power	Direct	6075	-8.35*	p < .001
	1Hop	10938	-4.85*	p < .001
	2Hop	24231	-5.23*	p < .001
Outlink Power	Direct	6075	-5.31*	p < .001
	1Hop	10938	-6.66*	p < .001
	2Hop	24231	-7.06*	p < .001
Overall Power	Direct	6075	-6.94*	p < .001
	1Hop	10938	-6.30*	p < .001
	2Hop	24231	-6.15*	p < .001

5.2.3 The number of connections and the number of resources

We found some evidence that information items may propagate along networks of trust. The number and the fraction of shared information items is largest for directly connected users, decrease with the increase of distance in the network and reach its minimum for non-connected users. If information, indeed, flows along the network of trust, we can hypothesize that users who have more trusted connections will be able to assemble more items in their own collection. To check whether this hypothesis is supported by our dataset, we correlated the number of resources that each user have and the number of trusted connection of this user. In *Citeulike* dataset, we discovered significantly positive relationships, $r = .214$, $p < .001$. However, we found no significant relationships between the sizes of user networks and collections in *Delicious* data set, $r = -.021$, $p = .162$.

Furthermore, the item popularity and how the popularity correlates with relationships were explored (H4). For this test, we select all items, which were stored by more than one user. In order to normalize the data distribution (the more users store items, the more pairs share the items), we computed the percentage of the unidirectional or reciprocal relations among the people who store the same information item.

While it is natural to expect that more popular items are shared by trusted users proportionally more frequently, the reverse appeared to be true. In *Citeulike* dataset, there was a significant negative relationship between the item popularity and the percentage of networks, $r = -.253$, $p < .001$. Namely, *less* popular items were shared *more* by trusted pairs, even the size of correlation was small ($r < .3$). This significant negative relationship was also found in the *Delicious* dataset, $r = -.087$, $p < .001$. Thus, more unique and rare information has a higher chance to be shared by trusted users. This fact reveals deeper-level interest similarity between users in networks of trust.

5.3 Tag Similarity in Trust Network

In this section, we examine whether interests of users connected by a relation of trust are also similar on the level of tags. Tag similarities were compared between users connected by direct and distant relations of trust and between non-connected users. As explained in Section 4, two kinds of tag similarities – micro level and macro level – were taken into consideration.

For both systems, users connected by a relation of trust shared more similar tags on both micro-level and macro-level (Tables 15 to 20). Both absolute numbers and relative fractions (inlink, outlink, and overall powers) of shared tags were the largest for users connected by direct relations of trust. The similarity decreased with the increase of distance reaching its minimum for non-connected users (infinite distance). For all conditions, the difference in similarity was significant between distances. Moreover, for all distances and both systems, tag similarity between users connected by a reciprocal relation was much larger than for users connected by a unidirectional relation. I.e., the similarity measured on the level of tags (both macro and micro) exhibited the same patterns as similarity on the level of items and sources.

Table 15. The Number of the Common Tags (Unidirectional Relation)

		Direct	1hop	2hop	No Rel.
Citeulike	Micro Level	**.28**	**.04**	**.02**	**.00**
		$F\ (3, 154540) = 667.568$, $p < .001$, $\eta^2 = .020$			
	Macro Level	**3.65**	**4.44**	**2.79**	**.29**
		$F\ (3, 154540) = 1627.306$, $p < .001$, $\eta^2 = .100$			
Delicious	Micro Level	**.16**	**.03**	**.02**	**.01**
		$F\ (3, 84125) = 284.826$, $p < .001$, $\eta^2 = .017$			
	Macro Level	**4.77**	**3.78**	**3.49**	**1.85**
		$F\ (3, 84125) = 824.326$, $p < .001$, $\eta^2 = .060$			

Table 16. The Number of the Common Tags (Reciprocal Relation)

		Direct	1hop	2hop
Citeulike	Micro	**4.60**	**.51**	**.18**

	Level	F (2, 1454) = 37.391, $p < .001$, $\eta^2 = .093$		
	Macro Level	**12.41**	**9.45**	**8.50**
		F (2, 1457) = 6.239, $p = .002$, $\eta^2 = .008$		
Delicious	Micro Level	**60.39**	**49.91**	**.02**
		F (2, 17105) = 8599.917, $p < .001$, $\eta^2 = .498$		
	Macro Level	**5.42**	**3.73**	**3.38**
		F (2, 17105) = 110.113, $p < .001$, $\eta^2 = .013$		

Table 17. Similarity Powers of Tags in *Citeulike* (Unidirectional Relations)

		Direct	1hop	2hop	No Rel.
Micro Level	Inlink	**0.39%**	**0.02%**	**0.01%**	**0.00%**
		F (3, 154540) = 336.041, $p < .001$, $\eta^2 = .010$			
	Outlink	**0.15%**	**0.04%**	**0.03%**	**0.00%**
		F (3, 154540) = 201.083, $p < .001$, $\eta^2 = .006$			
	Overall	**0.06%**	**0.01%**	**0.00%**	**0.00%**
		F (3, 154540) = 510.960, $p < .001$, $\eta^2 = .015$			
Macro Level	Inlink	**11.85%**	**5.18%**	**3.04%**	**8.33%**
		F (3, 154540) = 228.516, $p < .001$, $\eta^2 = .016$			
	Outlink	**6.02%**	**4.94%**	**5.05%**	**3.39%**
		F (3, 154540) = 277.480, $p < .001$, $\eta^2 = .019$			
	Overall	**2.17%**	**1.43%**	**1.07%**	**1.53%**
		F (3, 154540) = 184.511, $p < .001$, $\eta^2 = .012$			

Table 18. Similarity Powers of Tags in *Citeulike* (Reciprocal Relations)

		Direct	1hop	2hop
Micro Level	Inlink & Outlink	**1.28%**	**0.11%**	**0.04%**
		F (2, 1457) = 97.176, $p < .001$, $\eta^2 = .118$		
	Overall	**0.47%**	**0.03%**	**0.01%**
		F (2, 1457) = 156.765, $p < .001$, $\eta^2 = .177$		
Macro Level	Inlink & Outlink	**16.07%**	**9.47%**	**5.96%**
		F (2, 1457) = 56.820, $p < .001$, $\eta^2 = .072$		
	Overall	**5.43%**	**2.06%**	**1.31%**
		F (2, 1457) = 265.729, $p < .001$, $\eta^2 = .267$		

Table 19. Similarity Powers of Tags in *Delicious* (Unidirectional Relations)

		Direct	1hop	2hop	No Rel.
Micro Level	Inlink	**0.17%**	**0.04%**	**0.02%**	**0.01%**
		F (3, 84125) = 210.507, $p < .001$, $\eta^2 = .013$			
	Outlink	**0.22%**	**0.05%**	**0.02%**	**0.01%**
		F (3, 84125) = 170.669, $p < .001$, $\eta^2 = .011$			
	Overall	**0.08%**	**0.02%**	**0.01%**	**0.00%**
		F (3, 84125) = 273.044, $p < .001$, $\eta^2 = .016$			
Macro Level	Inlink	**8.86%**	**7.66%**	**7.00%**	**4.22%**
		F (3, 84125) = 661.734, $p < .001$, $\eta^2 = .050$			
	Outlink	**9.83%**	**7.16%**	**6.47%**	**3.66%**
		F (3, 84125) = 1106.841, $p < .001$, $\eta^2 = .082$			
	Overall	**3.83%**	**2.99%**	**2.75%**	**1.50%**
		F (3, 84125) = 1295.017, $p < .001$, $\eta^2 = .092$			

Table 20. Similarity Powers of Tags in *Delicious* (Reciprocal Relations)

		Direct	1hop	2hop
Micro Level	Inlink & Outlink	**63.21%**	**61.21%**	**0.03%**
		F (2, 17105) = 1269.757, $p < .001$, $\eta^2 = .129$		
	Overall	**18.78%**	**17.76%**	**0.01%**
		F (2, 17105) = 43951.64, $p < .001$, $\eta^2 = .835$		
Macro Level	Inlink & Outlink	**11.49%**	**7.74%**	**6.80%**
		F (2, 17105) = 171.214, $p < .001$, $\eta^2 = .020$		
	Overall	**4.61%**	**3.11%**	**2.65%**
		F (2, 17105) = 272.308, $p < .001$, $\eta^2 = .031$		

The final hypothesis was about whether popularities of tags were correlated with the number of cases where the corresponding tags were shared in trusted connections. We found the significant negative relationships in *Citeulike* dataset, $r = -.109$, $p = .042$. The tags shared within the trusted network were less popular. In *Delicious* dataset, we could not find any correlation, $r = .021$, $p = .672$.

6. CONCLUSION AND DISCUSSION

The paper argued that users' self-defined relations of trust could be valuable to increase the quality of recommendation in CF systems. To prove the feasibility of this idea we examined how similar are interests of users connected by a self-defined relation of trust. Interest similarity was measured by similarity of items and sources they share and tags they use.

Using *Delicious* and *Citeulike* datasets, we found that user connected by a self-defined relation of trust have more common information items, sources and tags than non-connected users. The similarity was largest for direct connections and decreased with the increase of distance between users in the network of trust. Users involved in a reciprocal relationship exhibited significantly larger similarity than users in a unidirectional relationship on all levels. Moreover, similarity on the level of sources (authors and web sites) was larger than similarity on the level of individual items (references and bookmarked pages). The information-sharing pattern were generally similar for *Citeulike* and *Delicious* although trusted pairs of users in a relatively specialized and homogeneous system *Citeulike* exhibited larger information similarity.

If trusted networks are a good source for providing similar information, is the information useful, as well? Granovetter [6] asked the same question in job search context. He compared the flow of information about job openings between weakly tied and strongly tied relations. He insisted that the influence of strongly connected people, such as friends, family, or relatives was not important regarding the dissemination of information. This is because the same information is circulated only inside of their network circle and hardly gets outside of the social network; therefore, it is difficult for new information to get inside of the network. Because of this, the information delivered from the weakly tied network is more useful than information from the strongly tied network.

The paper [6] also showed that the job opening information is much easier to obtain from an acquaintance than from people close to the individual. This trend grows stronger when the people have a higher education background and the information is more professional in nature [6]. As Grannovetter suggested through empirical analysis in another study, the strength of the relationship is proportional to their similarity [5]. In this knowledge-driven era, nevertheless, if a person wants to acquire new professional information it is more useful to take advantage of information available outside of their own limited social network [6]. This study implies two points worth pondering. Firstly, personal interactions are not essential elements in acquiring useful information. Weakly tied networks which have no or few personal contacts are quite common on the Web. Secondly, the question of how to compromise between information similarity and usefulness is arisen. To cope with questions about the usefulness of the shared information in a trusted network, in future, we will examine the recommendation quality, using our data sets. How the information propagates within the trusted network and the influence of information authorities who play a leading role in disseminating the information will both be investigated. In later studies, we plan to expand our target domains by adding different data sets.

7. REFERENCES

[1] Adamic, L. A. & Adar, E. (2003) Friends and neighbors on the Web, *Social Networks*, 25 (3), pp. 211 ~ 230.

[2] Ames, M. & Naaman, M. (2007) Why We Tag: Motivations for Annotation in Mobile and Online Media, In: *Proc. of CHI 2007* (San Jose, CA, USA, April 2007), pp. 971 ~ 980.

[3] Bonhard P. & Sasse M. A. (2006) Knowing me, Knowing you - Using Profiles and Social Networking to Improve Recommender Systems, *BT Technology Journal*, Vol. 25 (3)

[4] Firan, C. S., Nejdl, W. & Paiu, R. (2007) The Benefit of Using Tag-based Profiles, In: *Proceeding of the 17th ACM conference on Information and knowledge management*, Napa Valley, CA, USA, pp. 193-202

[5] Granovetter, M. (1973) The Strength of Weak Ties, *American Journal of Sociology*, 78 (6), pp. 1361 ~ 1380

[6] Granovetter, M. (1983) The Strength of Weak Ties: A Network Theory Revised, *Sociological Theory*, Vol.1, pp. 201 ~ 233.

[7] Hotho, A., Jäschke, R., Schmitz, C. & Stumme, G. (2006) Trend Detection in Folksonomies, *Semantic Multimedia*, pp. 56-70

[8] Hung, C., Huang, Y., Hsu, J. Y. & Wu, D. K. (2008) Tag-based User Profiling for Social Media Recommendation, In: *Proceedings of the Workshop on Intelligent Techniques for Web Personalization and Recommender Systems, the 23rd AAAI Conference on Artificial Intelligence (AAAI-08)*, Chicago, Illinois, USA, July 2008.

[9] Lam, S. K. & Riejl, J. (2004) Shilling Recommender Systems for Fun and Profit, In: *Proceedings of World Wide Web 2004*, New York, NY, USA, pp. 393 ~ 402

[10] Lund, B., Hammond, T., Flack, M. & Hannay, T. (2005) Social Bookmarking Tools (II): A Case Study – Connotea, *D-Lib Magazine*, 11(4).

[11] Massa, P. & Avesani, P. (2004) Trust-aware Collaborative Filtering for Recommender Systems, In: *Proceedings of Federated International Conference On The Move to Meaningful Internet*, Agia Napa, Cyprus, pp. 492 ~ 508

[12] Massa, P. & Avesani, P. (2007) Trust-aware Recommender System, In: *Proceedings of Recommender System 2007*, Minneapolis, MN, USA, pp. 17 ~ 24

[13] Maltz, D. & Ehrlich, K. (1995) Pointing the Way: Active Collaborative Filtering, In, Proceeding of CHI' 95, Denver, CO, USA, pp. 1 ~ 8

[14] McKnight, D. H. & Chervany, N. L. (1996). The Meanings of Trust. University of Minnesota, http://misrc.umn.edu/wpaper/WorkingPapers/9604.pdf (accessed on July, 2008)

[15] Mehta, B., Hofmann, T., Nejdl, W. (2007) Robust collaborative filtering, In: Proceedings *of Recommender System 2007*, Minneapolis, MN, USA, pp. 49 ~ 56

[16] Nov, O., Naaman, M. & Ye, C. (2008) What Drives Content Tagging: the Case of Photos on Flickr, In: *Proceedings of CHI 2008* (Florence, Italy, April 2008), pp.1097 ~ 1100.

[17] O'Donovan, J. & Barry, S. (2005) Trust in Recommender Systems, In: *Proceedings of the 10th International Conference on Intelligent User Interfaces*, San Diego, California, USA, pp. 167 ~ 174.

[18] Schafer, J. B., Frankowski, D., Herlocker, J. & Sen, S. (2007) Collaborative Filtering Recommender System, In: Brusilovsky, P., Kobsa, A. & Nejdl, W. (Eds.) *The Adaptive Web: Methods and Strategies of Web Personalization*, Springer, Berlin, Germany, pp. 291 ~ 324

[19] Schenkel, R., Crecelius, T., Kacimi, M. & Neumann, T. (2008) Social Wisdom for Search and Recommendation, *IEEE Data Engineering*, 31(2), pp. 40 ~ 49.

[20] Singla, P. & Richardson, M. (2008) Yes, There is a Correlation - From Social Networks to Personal Behavior on the Web, In: *Proceeding of the 17th International World Wide Web Conference*, Beijing, China.

[21] Sinha, R. & Swearingen, K. (2001) Comparing Recommendations Made by Online Systems and Friends, In: *Proceedings of DELOS Workshop on Personalisation and Recommender Systems*

[22] Tintarev, N. & Masthoff, J. (2007) Effective explanations of recommendations: user-centered design. In: *Proceedings of Recommender System 2007*, Minneapolis, MN, USA, pp. 153-156

[23] Zhao, S., Du, N., Nauerz, A., Zhang, X., Yuan, Q. & Fu, R. (2008) Improved Recommendation based on Collaborative Tagging Behaviors, In: *Proceedings of International Conference on Intelligent User Interfaces*, Maspalomas, Spain, pp. 413 ~ 416.

[24] Ziegler, C. & Golbeck, J. (2007) Investigating interactions of trust and interest similarity, *Decision Support Systems*, 43, pp. 460 ~ 475