Modeling Trust for Recommender Systems using Similarity Metrics

Georgios Pitsilis and Lindsay F. Marshall

Abstract. In this paper we present novel techniques for modeling trust relationships that can be used in recommender systems. Such environments exist with the voluntary collaboration of the community members who have as a common purpose the provision of accurate recommendations to each other. The performance of such systems can be enhanced if the potential trust between the members is properly exploited. This requires that trust relationships are appropriately established between them. Our model provides a link between the existing knowledge, expressed in similarity metrics, and beliefs which are required for establishing a trust community. Although we explore this challenge using an empirical approach, we attempt a comparison between the alternative candidate formulas with the aim of finding the optimal one. A statistical analysis of the evaluation results shows which one is the best. We also compare our new model with existing techniques that can be used for the same purpose.

1   Introduction

Recommender systems have become popular nowadays as they are widely used in e-commerce. Examples of services which use recommender systems for helping users to choose products they might like are epinions [1], eBay [2] and Amazon [3]. The contribution of recommender systems comes in two forms, either as predicted ratings of services that a user wants to know about, or as lists of services that users might find of interest. The effectiveness of a *Recommender system* can be measured by the accuracy of the predictions that it makes. *Collaborative filtering* (CF) [4] is

Georgios Pitsilis and Lindsay F. Marshall
School of Computing Science, University of Newcastle Upon-Tyne, Newcastle Upon Tyne, NE1 7RU, U.K., e-mail: Georgios.Pitsilis@ncl.ac.uk

the most widely known technique used in Recommender systems and is based on the idea of making predictions using similarity metrics to correlate users.

However, *Recommender Systems* and particularly *Collaborative Filtering* are not perfect and as it is well known that they appear to have weaknesses such as a low quality of predictions (known as the *false negatives* and *false positives* problems [5]), caused by sparsity in the dataset. Also, the architectural characteristics of CF are known to be vulnerable to attacks from malicious and libelous users. CF systems employ statistical techniques to develop virtual relationships between users, and in this way, neighborhoods of users can be formed consisting of those who have a history of agreeing and who are thus assumed to be similar. The virtual relationships are built upon a metric that is used for correlating the users based on their experiences and is called *Similarity*. In order to know how similar two users are with each other, a number of common experiences must exist.

Trust has been investigated by many researchers of recommender systems in the past [23] and proposed also as a potential solution to alleviate the previously mentioned problems of recommender systems [6,7]. Trust can also express integrity in relationships between entities and so can be used to express the quality of service providers. So, service consumers should be able to assess reliably the quality of services before they decide to depend on a particular instance. In order to know the trustworthiness of a service provider evidence needs to be provided to potential consumers from which they can derive their own trust for the provider.

Under appropriate circumstances (with regard to a common purpose), trust relationships can also support transitivity [8] whereas similarity generally does not. In order to benefit from the special characteristics of trust such as the ability to propagate along chains of trusted users, a formula for deriving it from similarity and vice versa is needed. In this way user entities that cannot be correlated due to lack of common experiences can benefit from each other and thus extend the quantity and/or the quality of predictions they can make about their future choices. Our contribution to this research problem is the provision of appropriate formulas that can be used for converting trust to similarity.

The rest of the paper is organized as follows. In the next section, there is a more detailed description of the problem. Section 3 includes related work in the field and in section 4 we analyze our approach to the problem, showing the formulas we have introduced. Next in section 5 we present the evaluation we performed and some comparative results which show the best candidate. Finally, in section 6 we discuss some future issues concerning the applicability of the proposed method.

2   Motivation

The main idea of collaborative filtering is to make predictions of scores based on the heuristic that two people who agreed (or disagreed) in the past will probably agree (disagree) again. A typical collaborative filtering system runs as a centralized service and the information it holds can be represented by a matrix of users and items. Each value of the matrix represents the score that a particular user has given to some item. The number of empty cells is known as sparsity and as we mentioned in the previous section, it is the main reason that recommender systems behave poorly, be-

cause not much evidence can be gathered to support a recommendation. This is usually because users themselves are unwilling to invest much time or effort in rating items. In existing CF systems users can only be correlated through their common experiences, so in the presence of limited data they turn out to be unable to make accurate predictions. The idea to enhance the neighboring base of users, by using the potentially developed trust relationships between them, could make it possible to reach other members of the community through them.

Assuming that the potential trust between the users could help in reducing the number of empty cells in the matrix by allowing missing values to be predicted from existing ones, finding a way of computing that trust from the existing data (user experiences) might help to alleviate the problem.

For such an idea to be applicable, it is necessary that, somehow, users must be able to place trust on their neighbors. In some centralized consumer opinion sites [1] it is a requirement that this trust measure should be provided by the users themselves. However, this requires that users should have developed some instinct in judging things accurately, and this cannot be assured. Poor judging abilities introduce the danger of establishing relationships with wrong counterparts. Our approach to this issue is to introduce a technique for mapping between similarity measures and trust, and which will be done automatically on behalf of the users.

In our model we use ordinary measures of similarity taken from CF to form the potential trust between the correlated entities which would be propagated in a similar way to the word-of-mouth scheme. In that scheme the trust that the first entity should place on the distant one is derived through a trust graph. Finally, by transforming the value back into similarity measure terms it could be made appropriate for use in CF algorithms. However, to our knowledge, today there is no standard approach for modeling trust from such type of existing evidence. In this work as well as in a previous one [9] we express trust in the form of opinions as they are modeled in *Subjective Logic* [10]. In this theory *trust* is considered as a subjective measure and introduces the important idea that there is always imperfect knowledge when judging things. The latter is expressed with a notion called *uncertainty* and is present when trust is based on user observations. Another interesting point of subjective logic is that it provides an algebra for combining *direct* and *indirect* trust along chains of users. *Direct* trust is considered the trust that is built upon first hand evidence or else derived from experience with the trustee. *Indirect* trust is built upon recommendations from others when first hand evidence is not present.

The use of trust in transitive chains requires the existence of a common purpose [8] which needs recommender trust to be derived or given from a specific transitive chain. This has either to be modeled from relevant evidence or, somehow, trustors must be enabled to derive it from past experiences.

Our work in this paper is concerned with the construction of trust relationships using first hand evidence, which in our case is the users' ratings. More specifically we try various similarity-to-trust transformation formulas with the purpose of finding the most suitable one. In the future we aim to evaluate the accuracy of a whole recommender system that employs the proposed transformation formula.

3   Background Research

Trust has long been a concern for scientists and much work has been done to formalize it in computing environments [11,12]. As well as being context specific, it has important characteristics such as asymmetry, subjectivity, and under specific circumstances, transitivity. It is also related to tasks in the sense that entities are trusted to perform a particular task. A simplistic approach would be to determine the levels of trust and distrust that should be placed on some entity from its probabilistic behavior as seen from trustor's point of view. In this sense, trust can be thought of as the level of belief established between two entities in relation to a certain context. In uncertain probabilities theory [13] the metric which expresses the belief is called *opinion*. Because there is always imperfect knowledge as opinions are based on observations, lack of knowledge should be considered when assessing them. *Subjective Logic* framework deals with the absence of both trust and distrust by introducing the *uncertainty* property in opinions. This framework uses a simple intuitive representation of uncertain probabilities by using a three dimensional metric that comprises belief (b), disbelief (d) and uncertainty (u). Between *b,d* and *u* the following equation holds b+d+u=1 which is known as the *Belief Function Additivity Theorem*. Building up opinions requires the existence of evidence, but even though opinions in the form (b,d,u) are better manageable due to the quite flexible calculus that opinion space provides, evidence is usually available only in other forms, that are essentially more understandable to humans.

Having this in mind, we could use the ratings given by the users as evidence, also called behavioral data, for forming trust relationships between them in a CF system. The *Beta Distribution Probability Function* can offer an alternative representation of uncertain probabilities [14], making it possible to approximate opinions from behavioral data. However, data in that evidence space are considered as sets of observations and therefore must be provided strictly in binary form representing the possible two outcomes of a process, x or $\overline{x}$. So, a behavior is described by the number of x and $\overline{x}$ that derives from the set of observations. In [10] there is a mapping between Evidence Spaces and Opinion Spaces where the uncertainty property (u) is solely dependent on the quantity of observations. In contrast, other similarity based approaches such as that in [15] are based on the idea of linking users indirectly using predictability measures, but, to our knowledge, these have not been tested in real environments.

As we mentioned above, the requirement for trust to become transitive in long chains is that a common purpose exists along the chain. According to this, only the last relationship should be concerned with trust for a certain purpose and all the other trust relationships in the chain should be with respect to the ability to recommend for the given purpose. The former is called *functional* trust and the latter *recommender* trust. It is worth mentioning the existence of other approaches to making recommender systems trust-enabled such as [16] where there is no distinction between functional and recommender trust. Also in some other solutions [17] that are used for predicting scores in recommender systems using webs of trust, the notion of trust is confused with similarity even though they are essentially different. Subjective logic provides a useful algebra for calculating trust in long chains of neighbors but it requires that opinions be expressed in (b,d,u) format which existing modeling tech-

niques are not suitable to handle. This is because existing solutions for encoding trust either deal with data in an unsuitable form (see [14] beta pdf) or do not provide links to similarity. In our opinion it is not appropriate for users to be asked to provide trust measures for others, mainly because this requires skills and adequate experience that not all users have.

4   Our Approaches

In general, trust models are used to enable the parties involved in a trust relationship to know how much reliance to place on each other. Our model aims to provide a method for estimating how much trust two entities can place in each other, given the similarities between them.

The problem that emerges when Trust is to be used in a recommender system is the fact that the entities involved usually provide their views in the form of ratings about items and not as their trust estimates about other entities. That means, to benefit from such model it is required that all user ratings be transformed into trust values. We are contributing to solving this issue by proposing and comparing various formulas for encoding direct trust. The first formula we propose in paragraph 4.1 has already been used in an experimental P2P recommender system which has been studied in [18]. The other new modeling approaches we propose are extensions of the same idea. The significant difference, though, between the existing and the new approaches is found in the way we model the uncertainty property. In all the new approaches we keep the main method of modeling uncertainty the same but we change the way that the remaining properties (belief and disbelief) are shaped.

4.1   The existing approach

Unlike the other modeling concepts we discussed above, such as beta pdf modeling, in our first approach we use both quantitative and qualitative criteria on the evidence to derive uncertainty. In order to achieve this, we consider the ratings that users have given to items as the behavioral data required for the composition of opinions. In order to capture this requirement in our model we assume that the level of trust that develops between every pair of entities is based on how similar they perceive each other's choices to be. We used the *Pearson* coefficient, as this is the best known and most suitable coefficient for this type of application. It can take values between -1 and 1 where two entities are considered as having higher similarity when their Pearson values are close to 1 and as completely dissimilar when the Pearson Coefficient is -1. A value of 0 would mean that there is no relationship between the two entities at all. Bearing in mind the idea that those entities whose ratings can be accurately predicted should be considered as trustworthy sources of information, the uncertainty in such relationships should be lower.

Thus, in this approach we have re-defined the perception of *Uncertainty* as the inability of some entity to make accurate predictions about the choices of the other

counterpart in the relationship. A low ability value should result from the existence of conflicting data and this should make the observer unable to fill in the uncertainty gap. When there are not enough observations to distinguish rating trends data might appear to be highly conflicting.

We propose the following formula to model uncertainty from prediction error:

$$u = \frac{1}{k} \sum_{x=1}^{k} \frac{|p_x - r_x|}{m} \tag{4.1}$$

where k is the number of common experiences (ratings) of the two entities that form a relationship, p_x is the predicted rating of item x calculated using some prediction calculation formula and r_x is the real rate that the entity has given to item x. m represents the maximum value that a rating can take and it is used here as a measure of rating. As can be seen, uncertainty is inversely proportional to the number of experiences. This agrees with the definition of uncertainty we presented in the previous section.

The logical reasoning for deriving formula (4.1) for Uncertainty is the following: Uncertainty is proportional to the prediction error for every user's single experience; therefore the numerator represents the absolute error between the predicted value (using a rating prediction formula) and the real (rated) value. The denominator m has been used for normalizing the error to the range 0-1. The summing symbol has been used to include all the experiences (k in number) of a particular user. Finally, the division by the total number of experiences (k) is done to get the average normalized error. In the sum we take every pair of common ratings and try to predict what the rate p would be. Therefore it is assumed that on every prediction calculation all but the real rating of the value that is to be predicted exist.

Unlike Beta mapping [14] where u tends to 0 as the number of experiences grows, in our model the trend remains quite uncertain because u is also dependent on the average prediction error. In the extreme case where there is high controversy in the data, u will reach a value close to 1, leaving a small space for belief and disbelief. Another interesting characteristic of our model is the asymmetry in the trust relationships produced, which adheres to the natural form of relationships since the levels of trust that two entities place on each other may not be necessarily the same.

As regards the other two properties b (belief) and d (disbelief), we set them up in such a way that they are dependent on the value of the Correlation Coefficient CC. We made the following two assumptions:
- The belief (disbelief) property reaches its maximum value (1-u) when CC=1 (or CC=-1 respectively)
- The belief (disbelief) property reaches its minimum value (1-u) when CC= -1 (or CC=1 respectively)

which are expressed by the two formulae:

$$b = \frac{(1-u)}{2}(1+CC) \tag{4.2}$$

$$d = \frac{(1-u)}{2}(1-CC) \tag{4.3}$$

As can be seen, the ratio of belief and disbelief is shaped by the CC value. In this way, a positive Correlation Coefficient would be expected to strengthen the belief property at the expense of disbelief. In the same way, disbelief appears to be stronger than belief between entities that are negatively correlated (CC<0).

These two formulae can be used in the opposite way too : for estimating how similar the two entities should consider each other, given their trust properties. The asymmetry in the trust relationships is mainly responsible for having unequal similarities between the original one and the one derived from the backward application of the formula. The different points of view are responsible for this difference as well as the formula used to work out the predictions p_x in (4.1). The formulas proposed in [15] as well as Resnick's [19] empirical one built for the Grouplens CF system can be used for the calculation of p_x.

As we can see in this proposed model, belief/disbelief increases/decreases linearly with the Correlation Coefficient and in terms of computational complexity, the uncertainty formula is $O(n^2)$. This seems to be a significant drawback to this method because the calculation of uncertainty requires the prediction formula to run for n times which in turn requires the calculation of similarity value k times. This has to be repeated whenever a new score is entered by any of the two parties.

4.2   The new proposed model

Since the above formula is found to be computationally intensive we came up with other less complex alternative formulas for modeling the same notions.

The first thing that we changed was the calculation of uncertainty. In contrast to the old approach, in the new design it is calculated exclusively from the quantity of experiences similarly as is done in the beta pdf mapping in Josang's approach [14]. However, in our new model we propose that every pair of common scores is counted as a different experience and for the uncertainty calculation we use the formula: $u = (n+1)^{-1}$, where n is the number of common scores.

As to *belief* and *disbelief* we tried various associations with CC such as linear, non-linear and circular. Amongst the pros of the alternative formulas is the significantly lower complexity $O(n)$ which means lower calculation time since it is now dependent only on the number of common ratings.

For a linear approach to shaping belief and disbelief the formulae used should be the same as before in the original model expressed in (4.2) and (4.3). For non-linear approaches we tried equations which are shown as figures of various skewnesses. The belief property alternatives are expressed in table 1. To save space, the formulas from which disbelief (*d*) is derived are not presented but for all cases *d* is considered as the remainder since d = 1 – b – u and it is symmetric to belief.

In addition to the two assumptions we made for the linear mapping shown in the previous paragraph, we included a third which is:

- A zero correlation coefficient (CC=0) should mean that belief equals disbelief.

Next, in Table 1 we present all formulas we came up with for shaping the *b* property and conform with the 3 assumptions we made.

1.
$$b = \frac{1}{2}\left(\sin(CC \cdot \frac{\pi}{2}) + 1\right) \cdot (1 - u)$$
(4.4)

2.
$$b = \frac{1}{2} + \left(\frac{\arcsin(CC)}{\pi}\right) \cdot (1 - u)$$
(4.5)

3.
$$b = \frac{1}{2}(1 - u)\left(1 + CC^{\frac{1}{K}}\right)$$
(4.6)

4.
$$b = \frac{1}{2}(1 - u)\left(1 + CC^{K}\right)$$
(4.7)

Table 1..The proposed formulas for belief property.

Fig. 1 shows the form of all formulas used for shaping the belief property presented in Table 1, type 1 and 2 as well as for types 3 and 4 for various skewness *k*. The linear approach that we described in paragraph 4.1 is also shown in Fig.1
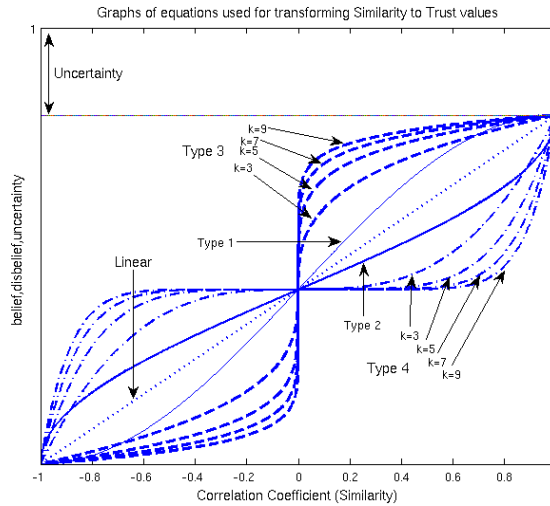

Fig. 1. The graphs of belief property for all formulas

5   Evaluation

When carrying out this experiment we faced the challenge of how to evaluate every alternative formula and what measures to use for comparing the accuracy of our modeling approach. Therefore, we developed and applied the following plan.

5.1   The plan

Since the goal was to test the accuracy of each candidate formula we considered as the best scenario comparing a known and accepted value of similarity against one that is derived by applying our trust derivation mechanism. More specifically, for each pair of users, lets call them A and B, we first calculated how similar they are, applying Pearson's CC formula over the common experiences of A and B, and then we calculated the indirect trust between them. Next, this trust value was converted to a similarity metric using our formula and, finally, the derived value was compared against the original similarity we calculated first. The latter similarity is derived from the resulting indirect trust between A and B when subjective logic rules are applied to the graph built by the trust relationships that exist between A and B. (see figure 2.) In order to accomplish this, the primary trust between every pair of users has to be built pro-actively when making up the trust graph.

Figure 2 is a pictorial representation of the entities involved in the evaluation scheme. We call $S_{A,B}$ the similarity that is derived from the common experiences between A and B, and $S'_{A,B}$ the similarity that is derived from the indirect trust of A for B. In the evaluation we compare these two values and we calculate the mean error.
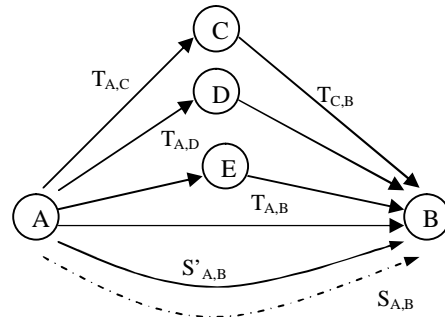

Fig. 2. . The evaluation diagram.

$T_{A,C}$ and $T_{A,D}$ are two of the direct trust values that are used for calculating the indirect (or secondary) trust of A for B.

Due to the fact that Pearson's coefficient has unstable behavior when there is a low number of common experiences between two parties, we considered as similar neighbors those who have at least 10 common experiences and we choose to perform the evaluation test on these pairs of entities as Pearson's similarity is calculable.

To measure the accuracy we calculated the Mean Absolute Error between the directly calculated similarity S and the one derived from the transitive trust S'. We use the following formula in which C_{max} and C_{min} are the maximum/minimum values of the Correlation Coefficient (1 and -1 respectively):

$$MAE = \frac{|S - S'|}{C_{max} - C_{min}} \tag{5.1}$$

The evaluation algorithm can be described in pseudo-code as in fig. 3. Let us call $dt_{i,j}$ the direct trust between entities i and j and $it_{i,j}$ the indirect one. Assuming that j is within 2 hops of i in the constructed trust graph, the indirect trust of i for j can be calculated using subjective logic in two steps: First, the derived trust of every alternative path that begins from i and ends to j is calculated separately as a transitive relationship using the suggestion operator $\otimes$. Then all the values of the alternative paths along with $dT_{i,j}$ are combined together using the consensus operator $\oplus$ which gives the value of $iT_{i,j}$. In general the consensus is expressed in the following formula where A and B are two different agents which hold about the statement p respectively the opinions ω_n^A and ω_n^B.

$$
\begin{array}{ll}
\underline{\text{Let } K} \text{ be the set of all users} \\
\underline{\text{Let } R} \text{ be the set of all ratings over items} \\
\underline{\text{Let } R_u} \subset R \text{ be the set of the ratings of some user } u \\
\underline{\text{Let } Ki} \subset K : |R_u| \geq 10 & \quad * \quad \text{Cardinality of set of ratings of user } i \; * \\
\underline{\text{For } i} \text{ in } Ki \\
\quad \underline{\text{Let } E_i} \subset R & \quad * \quad \text{The set of ratings of user } i \quad * \\
\quad \underline{\text{Let } M} \subset Ki : \forall p \in M, E_p \subset R \text{ and } |E_i \cap E_p| \geq 10 \\
\quad \underline{\text{For } j} \text{ in } M \text{ do} & \quad * \quad p \text{ has 10 common ratings with } i \\
\quad\quad S \leftarrow CC(i, j) & \quad * \quad \text{Pearson's similarity } * \\
\quad\quad T \leftarrow iTrust(i, j) & \quad * \quad \text{Derived } Indirect \text{ trust } * \\
\quad\quad S' \leftarrow f(T) & \quad * \quad \text{Derived Similarity from our formula } f \quad * \\
\quad\quad MAE \leftarrow \frac{|S - S'|}{C_{max} - C_{min}} & \quad * \quad \text{Absolute Mean Error value } * \\
\quad \underline{\text{End For } j} \\
\underline{\text{End For } i} \\
\underline{Average}(MAE)
\end{array}
$$

Fig. 3. The evaluation algorithm

The consensus opinion held by an imaginary agent A,B is:

$$\omega_p^{A,B} = \omega_p^A \oplus \omega_p^B = \{b_p^{A,B}, d_p^{A,B}, u_p^{A,B}\} \tag{5.2}$$

More about this can be found in [20]. In our particular case the statement p is the trustworthiness of the target j. A, B represent the alternative paths from i to j.

The algorithm for calculating the indirect trust between the origin *i* and the target *j* is shown in figure 4.

$$it_{i,j} \leftarrow dt_{i,l} \qquad\qquad\qquad\qquad \text{* direct trust is considered   *}$$

<u>Let</u> $E_i \subset R$ * the set of ratings of i *

<u>Let</u> $P \subset K_i : \forall f \in P, E_f \subset R$ and $\left| E_i \cap E_f \right| \geq 10$

<u>For</u> j in P do   * i has 10 common ratings with f *

      $t_{l,j} \leftarrow dt_{l,j}$

      $t_{i,l,j} \leftarrow t_{i,l} \otimes t_{l,j}$ * Serial graph composition *

      $it_{i,j} \leftarrow it_{i,j} \oplus t_{i,l,j}$ * Parallel graph composition *

<u>End For</u> j

<u>Return</u> $it_{i,j}$ as *iTrust*   * return the indirect trust *

Fig. 4. The indirect trust calculation.

The choice for exploring the trust graphs up to maximum distance of 2 hops was made mostly for reasons of simplicity since with a third hop the number of required calculations increases significantly without a corresponding substantial gain in accuracy.

Assuming that the trust transitivity mechanism of Subjective logic is accurate enough then any error measured via our experiment should be considered as error derived from our Similarity-to-Trust transformation formula. In order to evaluate our modeling approach we needed a suitable dataset of user's scores. We chose a publicly available dataset taken from a real CF system known as *MovieLens* [21]. MovieLens is a movie recommendation system based on collaborative filtering established at the University of Minnesota. The available dataset contains 1.000.209 anonymous ratings of approximately 3.900 movies made by 6.040 users who joined the service over the year 2000. For our experiment we used a subset of MovieLens that comprised 130000 ratings which were given by 1000 randomly selected users.

5.2   *Comparative Results and Discussion*

In our tests, we applied each candidate formula to 10 different data sets and the results were averaged. Each 100 user dataset built trust graphs of approximately 5000 relationships.

First, for each pair of users for which indirect trust is calculable the real and the trust-derived similarity are compared. The results are shown in figure 5. Also interesting to see is the measured correlation between the number of common experiences of *i* and *j* and the calculated error. Figure 6.a shows the above result pictorially including each pair of trusted entities.
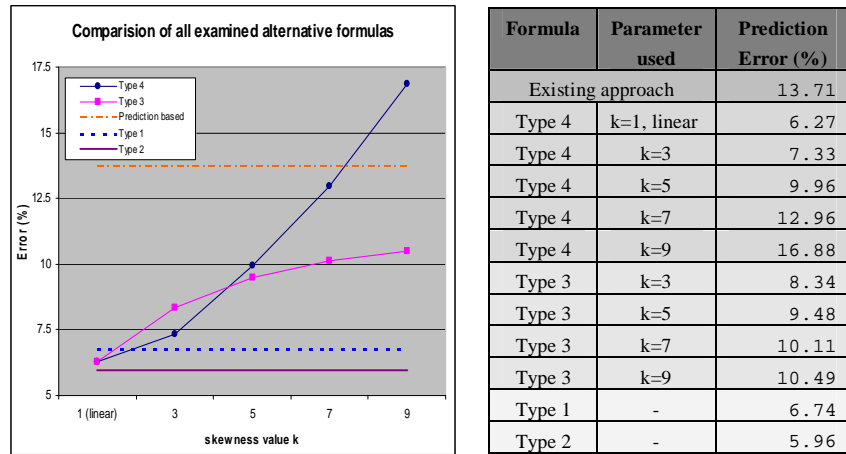
Formula	Parameter used	Prediction Error (%)
Existing approach		13.71
Type 4	k=1, linear	6.27
Type 4	k=3	7.33
Type 4	k=5	9.96
Type 4	k=7	12.96
Type 4	k=9	16.88
Type 3	k=3	8.34
Type 3	k=5	9.48
Type 3	k=7	10.11
Type 3	k=9	10.49
Type 1	-	6.74
Type 2	-	5.96

Fig. 5. Comparison of the alternative formulas.

Due to space limitations we present only the last candidate formula we tested (Type 2 of table 1) and which appears to give the best results (lowest error) compared to all other candidates. As can be seen from figure 6.a there is a stochastic relationship between the error of the transformation formula applied on pairs of users and the number of common scores of those users. It can also be seen that the maximum error observed is just above 35% as opposed to our first (*Existing*) approach which produced max error 70%. In terms of average error the new formula is better than twice as good. Another interesting, and obvious, observation from the figure is that as the number of common scores increases the error follows the opposite trend. Also, the deviation of the error decreases as the common scores increase. The importance of this observation is that it may have a practical value since it makes it possible to predict how accurately the derived trust will be calculated. Thus, when a recommendation is to be created a decision can be formed about whether or not a particular relationship should be considered in the process of secondary trust calculation. Hence, according to some quality criteria that can be applied, such primary trust relationships that have been built upon 'poor' data can be disclosed as they do not provide adequate contribution in the secondary trust calculation.

In figure 6.b we also show the variance of the similarity prediction error which has the practical meaning of how accurately the trust could be approximated using our model. That could have a practical use: for example it might be used as a criterion for choosing the right threshold value for the minimum number of common scores. In this way, the expected error can be determined according to some quality of service criteria that need to be met when building a web of trust.
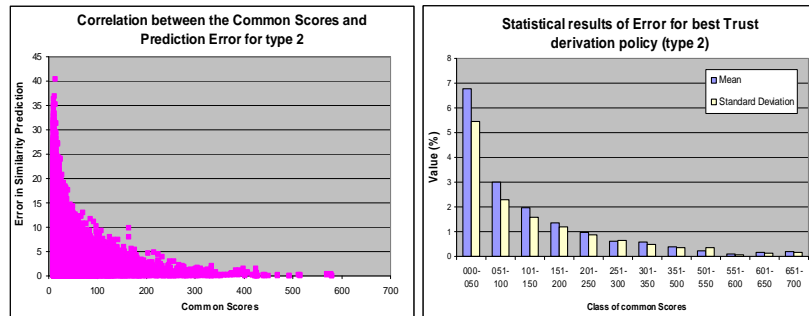
Fig. 6a (left) and **6.b.** (right). Graphical presentation of results.

We also looked to see if there is any correlation between the error of our trust de-rivation method and the number of common experiences between the two parties and the results show that there is slight correlation when considering linear approx-imation between error and common experiences. More specifically the correlation value declines as the *k* factor which expresses the skewness (see fig.1) increases. A more detailed study revealed that the error adapts best to a non-linear approxima-tion. The regression analysis on the results presented in fig.6.a. and are referred to the type 2 equation showed that the best value of Coefficient of Determination (R^2) for the above data had a relatively low value of -0.4135.

The increased divergence that is observed for the error as the number of common scores declines can be justified as the result of the noisy behavior of the Correlation Coefficient; therefore the quality of predictions is quite uncertain. Finally, it worth noticing that we observed higher prediction error when the correlation coefficient was between 0 and -1. This can be interpreted as: prediction is easier when users tastes agree and vice versa, or else the proposed formula is not as useful for disa-greeing tastes as it is when users agree. This means that using a unique mapping formula for the whole range of correlation coefficient values is not an ideal solution. Otherwise, the first assumption we have made in which, a similarity value of zero would mean that the trustworthiness with the other party should be the half of which corresponds to similarity value 1, is not absolutely right.

6   Future Work

As shown from a more careful examination of the results, there is high variation in the error in a way that follows different trends as *k* changes. For example the error that is measured between a certain pair of nodes *i* and *j* does not follow the same trend as that of the average. As also shown in figure 6.b and as discussed in the pre-vious paragraph, the varied deviations in the error need a more detailed analysis of the results to see if it will be possible to justify this observation.

For these two reasons it is worth investigating if and how the topological proper-ties of the derived network might be responsible for the variation and if it might be possible to decrease the error even more.

The assumption we have made in our modeling that the value of trustworthiness should correspond to a similarity value of zero seems to be not an optimum choice, therefore more investigation is needed towards finding the optimum fitting. As mentioned in paragraph 4.1 and 4.2 about the complexity of the formulas used, there is a potential problem since these computations need to be repeated for every user in the system. So, whenever a new user joins the system the trust and similarity computations will have to be redone against all existing users. As the system grows the computation time will increase significantly raising a scalability issue. A possible solution to this problem is to restrict the correlation process to a subset of participants rather the whole world. A focus of a future research is to investigate if clustering [22], a technique that is used for tackling a similar problem in recommender systems, can be quite effective here.

We intend to apply our technique to a real recommender system, with the expectation that it will improve the quality of the derived recommendations. Another idea is to make use of the web-of-trust that could evolve from the establishment of direct trust relationships between users. Our aim is to improve recommendations by exploiting the experiences of any entities not neighboring the querying one but which can be reached via the web-of-trust. The question that arises from this is how accurate these predictions can be. Short tests we performed, showed a significant increase in the coverage, which translates into reduced sparsity, without significant impact on the error in predictions. Our short-term plans include a thorough study and analysis of the various parameters that may have some impact on the results as well as a performance analysis of the resulting system. The long-term plans include the deployment of a totally distributed recommender system.

Recommender trust can be derived in a similar way to that described for *functional* trust in this paper. The basic idea is that someone's (lets call it the trustee) recommender trust can be estimated by some other entity (lets call it the trustor) by comparing any recommendations that trustee has provided in the past about statements for which the trustor also maintains its own evidence. Then the trustor, by comparing its relevant personal experiences with the trustee's recommendations, will be able to estimate how good in doing recommendations the trustee has been. Similarly to *direct* trust, *recommender* trust is a subjective measure, which means, every trustor has to maintain its own picture of the community.

No matter the quality of the recommendations such architecture can provide, there are weaknesses concerning security for the recommender systems that must also be covered. In particular, any deployed solution must be resistant to attacks from users that try maliciously to influence the system.

7   Conclusion

We presented an empirical technique for modeling the trustworthiness of entities using evidence that describe their rating behavior. The novelty comes from way that trustworthiness is derived from Similarity using a non-linear mapping.

We coded our derived trust opinions into metrics taken from Shaferian belief theory and we attempted an evaluation of our model by comparing the resulting similarity to that derived from secondary trust.

We also compared the proposed approach against another that has been used in the past and in which the shaping of the derived uncertainty is dependent on a predictability measure and thus on the quality of the evidence. The comparison showed that using qualitative measures for deriving trust not only incur a computation penalty but also provide lower accuracy when compared with less complex approaches for describing the user's behavior.

In conclusion, the strong points of the proposed technique can be summarized as its ability to incorporate similarity measures in its properties, the low computation complexity and its flexibility in accepting datasets in which user ratings are expressed in continuous values. In terms of accuracy in deriving trust opinions, a comparison against the older alternative shows that the new one is more than twice as accurate. We would suggest that the method is very suitable for use in CF recommender systems.

References

1. http://www.epinions.com
2. http://www.ebay.com
3. http://www.amazon.com
4. D.Goldberg, D.Nichols, B.M.Oki, D.Terry, "Using Collaborative filtering to weave an information tapestry", Communication of ACM, 35(12):61-70, 1992
5. B.Sarwar, G.Karypis, J.Konstan, J.Reidl, "Analysis of Recommendation Algorithms for ECommerce", In Proceedings of the Second ACM Conference on Electronic Commerce pg 158-168,ACM Press 2000.
6. M.P. O'Mahony , N.J. Hurley , G.C.M. Silvestre, "Detecting noise in recommender system databases", Proceedings of the 11th international conference on Intelligent user interfaces, January 29-February 01, 2006, Sydney, Australia
7. B. M.Sarwar , J.T.Riedl, "Sparsity, scalability, and distribution in recommender systems," 2001, ISBN:0-493-04207-5, University of Minnesota.
8. A. Jøsang, E. Gray and M. Kinateder. "Analysing Topologies of Transitive Trust", In the proceedings of the Workshop of Formal Aspects of Security and Trust (FAST) 2003, Pisa, September 2003.
9. G.Pitsilis., L.F.Marshall, "A model for trust derivation from evidence for use in recommender systems", Newcastle University, Technical report CS-TR-874, 2004.
10. A.Josang, "A Logic for Uncertain probabilities", International Journal of Uncertainty, fuzziness and Knowledge based systems, Vol.9,No.3, June 2001.
11. S.Marsh, "Formalizing Trust as Computational concept", PhD Thesis, University of Stirling, Scotland 1994.
12. A.Rahman, S.Heiles,"Supporting trust in Virtual Communities", In proceedings of International conference On System Sciences, Jan 4-7-2000, Hawaii.
13. G.Shafer, "A Mathematical Theory of Evidence", Princeton University Press. 1976
14. A. Jøsang, R. Ismail. "The Beta Reputation System". In the proceedings of the 15th Bled Conference on Electronic Commerce, Bled, Slovenia, 17-19 June 2002.
15. C.C.Aggarwal, J.L.Wolf, K.Wu, P.S.Yu,."Horting Hatches an Egg: A New Graph-theoretic Approach to Collaborative Filtering", In Proceedings of the ACM KDD'99 Conference. San Diego, CA, pp.201-212.
16. P.Massa – P.Avesani, "Trust-aware Collaborative Filtering for recommender Systems", CoopIS/DOA/ODBASE (1) 2004: 492-508

17. M.Papagelis, D.Plexousakis, T.Kutsuras, "Alleviating the Sparsity Problem of Collaborative Filtering using Trust Inferences", In Proceedings of iTrust, pp.224-239, 2005.

18. G.Pitsilis, L.F.Marshal, "Trust as a Key to Improving Recommendation Systems". In Proceedings of iTrust, Paris - France : pp.210-223, 2005

19. P.Resnick, N.Iacovou, M.Suchak, P.Bergstrom, J.Riedl, "Grouplens. An OpenArchitecture for Collaborative filtering of Netnews" In Proceedings of ACM 1994, Conf. On Computer Supported Cooperative.

20. A. Jøsang, "An Algebra for Assessing Trust in Certification Chains", In Proceedings of the Network and Distributed Systems Security (NDSS'99) Symposium, Internet Society, 1999

21. B.N.Miller,I.Albert, S.K.Lam, J.A. Konstan, J.Riedl, "MovieLens Unplugged: Experiences with an Occasionally Connected Recommender System",.In Proceedings of ACM 2003 International Conference on Intelligent User Interfaces (IUI'03) (Poster), January 2003.

22. B.M.Sarwar, G.Karypis, J.Konstan, J.Riedl, "Recommender Systems for Large-Scale E-Commerce: Scalable Neighborhood Formation Using Clustering",In Proceedings of the Fifth International Conference on Computer and Information Technology, December 2002, East West University, Bangladesh.

23. J. O'Donovan , B. Smyth,"Trust in recommender systems", Proceedings of the 10th international conference on Intelligent user interfaces, January 10-13, 2005, San Diego, USA