Development of a reliable trust management model in social internet of things

Mrutyunjaya Panda*

Department of ECE,
Gandhi Institute for Technological Advancement (GITA),
Bhubaneswar-752054, India
E-mail: mrutyunjaya74@gmail.com
*Corresponding author

Ajith Abraham

Machine Intelligence Research Labs (MIR Labs),
Scientific Network for Innovation and Research Excellence,
P.O. Box 2259,
Auburn, Washington 98071, USA
and
IT4 Innovations – Center of Excellence,
VSB – Technical University of Ostrava,
17, Listopadu 15/2172,
Ostrava – Poruba, Czech Republic
E-mail: ajith.abraham@ieee.org

Abstract: In this paper, we introduce the first application of the belief propagation algorithm in the design and evaluation of trust management systems with an introduction of a novel paradigm of social internet of things (SIoT), as 'social network of intelligent objects', that are based on the notion of social relationships among objects. Further, we address a trust-making process, where a person needs to make a judgement about the trustworthiness of another community member where they do not have any prior knowledge about each other. Our proposed model uses various performance measures such as: centrality and transitivity measures for SIoT analysis and then employs hybrid fuzzy nearest neighbour with Bayesian belief network and Bayesian belief network with conditional independence to represent a trust-based evaluation. Bayesian belief propagation technique is used here to infer trustworthiness in a social context. Finally, we perform non-parametric Friedman two tail test for statistical significance of the results obtained for various approaches. The evaluation of the model is done on datasets collected from epinion.com and slashdog.org shows promising results that enable us to steer the interaction among the billions of objects which will crowd the future IoT.

Keywords: internet of things; IoT; trust management; trust computing; communication; fuzzy; belief network; transitivity; coverage; accuracy; non-parametric test.

Reference to this paper should be made as follows: Panda, M. and Abraham, A. (xxxx) 'Development of a reliable trust management model in social internet of things', *Int. J. Trust Management in Computing and Communications*, Vol. X, No. Y, pp.xxx–xxx.

Biographical notes: Mrutyunjaya Panda received his PhD in Computer Science from Berhampur University. He received his Master in Communication System Engineering from Sambalpur University. He is having 16 years of teaching and research experience. He is presently working as Professor and Head at Department of ECE, Gandhi Institute for Technological Advancement (GITA), Bhubaneswar, Odisha, India. He has published about 53 papers in international and national journals and conferences. He has also published five book chapters, two-edited books, and two text books to his credit. He is a programme committee member of various international conferences. He is acting as an editorial member of various international journals. His active area of research includes data mining, intrusion detection and prevention, social networking, mobile communication, wireless sensor networks, granular computing, cognitive science, internet of things, etc.

Ajith Abraham received his MS from Nanyang Technological University, Singapore, and PhD from Monash University, Melbourne, Australia. His research experience includes more than 24 years in the industry and academia. He works in a multidisciplinary environment involving machine (network) intelligence, cyber security, sensor networks, web intelligence, scheduling and data mining. He has given more than 70 conference plenary lectures/tutorials and invited seminars around the globe. He is an author/co-author of numerous publications and also received several citations (h-index = 65+). He serves/has served the editorial board of over 50 international journals, guest-edited 40 special issues on various topics. He is the Chair of IEEE SMC Society, Technical Committee on Soft Computing.

1 Introduction

In today's connected world, the interaction of people who are unknown to each other and have no knowledge of them is possible and sometimes common. At any time the stability of a society relies on the trust or no trust between its people. On the other hand, today, we face with information overload problem which causes uncertainty and risk. Dealing with this problem is possible only through reliance to other in a community. Trust management has attracted many researchers in the fields of computing science including soft security, computer networks, e-commerce, game theory, social networks, etc. The term 'trust' is relatively confusing, lack of coherence and sometimes contradictory, with a variety of meaning as found from a good number of literature on trust (McKnight and Chervany, 2001; Hussain and Chang, 2007), demand a complete formal unambiguous definition of trust. However, the most cited definition of trust is found in Dasgupta (1990), where the author define trust as "the expectation of one person about the actions of others that affects the first person's choice, when an action must be taken before the actions of others are known". According to Golbeck (2006) trust can be defined as a person's commitment to an action based on belief that the future actions of that person will lead to a good outcome. This definition has a great limitation that it considers trust as always leading to positive outcome. But in reality, negative outcome is also possible. Considering the diversified concept and cross disciplines or domains, the definition of trust differs on the basis of the goal and the scope of the projects. Jøsang et al. (2007) provided two generalised definitions of trust, i.e., reliability trust (or evaluation trust) and

decision trust where evaluation trust can be interpreted as the reliability of something or somebody and the decision trust captures broader concept of trust.

In IoT, trusted devices are only the authorised object to access resources. The access credentials can be exchanged, and evaluated mechanically using trust negotiation. Binding trust and identity together addresses important issues like privacy protection, identity theft. Using efficient trust model, scalability can be achieved which is the one of the most important design issues in the context of IoT. Adequate management of identities in IoT is crucial to provide security and access control.

The problem statement for this research is the lack of trust in social network as internet usage progresses that plays a vital role in poor recognition as friend or foe with a higher risk. This problem may be identified or made evident through the behaviour of one over the other that generates a lack of trust and respect between the social network users, creates a feelings of disassociation among them.

The ultimate purpose of this social networking using internet of things (IoT) research is to identify the importance of trust from the perspective of social network users and to identify best practices for improving trust-building efforts within the community of interest using some novel evolutionary algorithms. It is to build relationships upon trust and respect.

To that end, this research is expected to produce findings and recommendations of best practices for building a trust model in social internet of things (SIoT) using Epinion and Slashdot Zoo datasets.

In this paper, we propose a data mining model for computing trust in Epinion and Slashdot Zoo dataset. This model computes the local trust for any person.

1.1 Solution to the problem statement

In this paper, we explore the utility of our trust management model by applying it to real datasets representing Epinion and Slashdot Zoo social networks. We first describe how to capture different types of interactions, recommend different things to different members in the community, and identify different roles in the community. Then using the above two datasets, we will demonstrate the efficacy of the model being developed. Finally, we address the question of sustainability with some novel data mining techniques in terms of accuracy, compute their performance based on transitivity, coverage, accuracy, mean absolute error (MAE) and F-score.

1.2 Solution to reliable trust model

Here, a reliable trust management approach in SIoT is thought of which can solve sparsity problems occurring in social trust models where actors are grouped using the similarity of attribute properties. In order to solve sparsity problems, we use fuzzy logic and Bayesian belief network (BBN) combined with conditional independence (CI) and then virtual relations are created between them in these groups to enrich relations. Finally, a reliable social trust model is presented by specifying their relationships with actual relations. The proposed approach is divided into an actual data extraction stage, a feature selection stage and then a trust management stage using data mining methodologies**.** In this paper, Epinion and Slashdot Zoo datasets are used to establish social networks.

Unfortunately, most of the existing approaches have been developed independently and little efforts have been made to compare and understand the relative strength and inherent vulnerabilities of different approaches. The novelty of this research lies in developing two-mode network visualisation of two most popular real social network datasets, application of fuzzy logic and BBN with CI for developing a probabilistic trust model. After this, performance evaluation is done with accuracy, similarity matrix, trust matrix, trust transitivity, MAE (maximum absolute error) and coverage, for understanding the proposed SIoT model. Here, we try to infer trust from distrust and care has been taken not to ignore any information.

The roadmap of the paper is organised as follows. Section 2 presents the related literature search. Social IoT and trust models are discussed in Section 3. In Section 4 we discuss and present the proposed evolutionary algorithms to develop the reliable trust model. Section 5 discusses about the experimental dataset (Epinion community, epinions.com) and Slashdot Zoo network dataset (from Slashdot.org) followed by the algorithmic steps and the experimental phases in Section 6 with a discussion on the perspectives of this work. Then, Section 7 concludes this report.

2 Related work

Trustworthiness management in the SIoT as a first theoretical analysis is discussed by Nitti et al. (2013) where the author provide an analysis of the performance of the subjective model, whose objective is to discriminate benevolent nodes from malicious ones with the minim error with subjective centrality measures. Robinson et al. (2010) presents trust and IoT where trust is used to control the presentation of the patterns and anchors within the augmented reality, building upon trust relationships that are dynamically created and maintained between the users of the system, yet maintain privacy. The authors (Sillence et al., 2006; Laia et al., 2011; Wua et al., 2010) address the factors of trust in specific domains considering the mapping between the evidence space and the trust space for evaluating the direct trust value based on the observations and evidences from the target behaviour. Computation of trust is measured using Dempster-Shafer theory in Qiu et al. (2010), where trust opinions are represented as mass assignments and then are combined with obtained rule of combination for the aggregated opinion. In Liao et al. (2011), Huynh et al. (2006) and Zhang and Zhang (2005), the author propose some kinds of trust models considering different issues related to trust management. SecuredTrust (Das and Islam, 2012) and CRM (Khosravifar et al., 2012) have addressed the concept of disposition to trust which is the inherent propensity of an individual to trust or distrust others along with its importance in trust management thereafter. A flexible framework for probabilistic models of social trust (Huang et al., 2013) with a soft-logic representation, demonstrates the flexibility and effectiveness of PSL for trust prediction on real social network data.

The author (Massa and Avesani, 2004) concluded that by incorporating trust, recommender systems can be more effective than systems based on traditional techniques like collaborative filtering. Further, they proposed that a peer can establish trust on other peers through explicit trust statements and trust propagation. It is worth noting that a trust model is built directly from users' direct feedbacks and then is incorporated into the recommendation process for recommending various items (such as books movie, music, software, etc.) to on-line users. Users can express their personal web of trust by

identifying those reviewers whose reviews and ratings are consistently found to be valuable. They pointed out that even though it is unclear how a user quantifies the degrees of trust when making trust statements, still it is possible to predict trust in unknown users by propagating trust with no direct connection between them. Similarity is measured using Pearson correlation coefficient on user-item ratings. In Massa and Bhattacharjee (2004), the author builds a trust model directly from trust data provided by users as part of the popular epinions.com service. The drawback of the work in Massa and Avesani (2004) and Massa and Bhattacharjee (2004) is that the web of trust are built on binary relationships among users and the propagating trusts are computed simply based on the distances between them. Massa and Avesani (2006) analyse the relative benefits of asking new users either few ratings about items or few trust statements about other users to generate recommendations. The experiments are conducted on a large real world dataset derived from Epinions.com with a conclusion that while traditional RS algorithms exploiting ratings on items fail for new users, asking few trust statements to a new user is instead a very effective strategy able to quickly let the RS generate many accurate items recommendations.

Now-a-days, the application of machine learning and soft-computing approaches for predicting the trust of service-oriented environments are suggested to be very successful in the literature. The authors (Wang et al., 2009) propose several trust evaluation metrics and a formula for trust computation, with which a final trust value is computed using fuzzy logic and using neural networks, decision trees, support vector machines-based classifiers are suggested for predicting the ordinal trust based on the QoS data (Al-Masri and Mahmoud, 2009; Mohanty et al., 2010).

The problem of managing trust in an open and centralised/decentralised system has attracted substantial research efforts in recent years (Caverlee et al., 2008; Hong and Shen, 2008; Kazai and Milic-Frayling, 2008; Golbeck, 2006; Santos-Neto et al., 2007). A commonly used solution to tackle the problem of trust management is to build a 'web of trust' where one helps the other in deciding whom to trust or to distrust, without prior interaction (Guha et al., 2004).

The authors (Shakeri and Bafghi, 2014) introduced a layer model of a confidence-aware trust management system with an abstract view of the main components of the system, their functions, and the relations among them, giving a global view of the system, facilitating modular engineering, simplifying the interoperability between trust researchers, and flexibility in evolution of the system. Idea towards a formalism that tries to describe trust in the IoT is discussed by the authors using a priori and a posteriori and explore the different meanings of trust and strategies that can be used to determine if something is trustworthy and propose a model for trust that takes into account people, devices, and their connections (Leister and Schulz, 2012).

There are many trust inference algorithms that include Advogato (Aiken and Levien, 1998), Appleseed (Ziegler and Lausen, 2004), Sunny (Kuter and Goldbeck, 2010), and Moletrust (Avesani et al., 2005) take advantage of pair-wise trust values and the structure of a social network. These algorithms use trust that is assigned on a fixed scale (e.g., 1–10). Other algorithms treat direct trust as a probability, including (DuBois et al., 2009; Hang et al., 2008; Patel et al., 2005; Josang et al., 2006). The difficulty of generating these probabilities, using influence as a proxy for trust, was addressed in Goyal et al. (2010). The authors (Quercia et al., 2006) propose a distributed trust framework that satisfies a broader range of properties.

In this paper, we focus on the problem of creating a framework for the trust inference, able to infer the trust/distrust relationships in those relational environments that cannot be described by using the classical social balance theory. Our framework evolves trust based on a Bayesian formalisation, whose trust metric is expressive to foster collaboration in a hostile pervasive computing environment. We investigate the evolutionary algorithms for predicting the trust with high level of accuracy.

3 IoT and trust models

3.1 Internet of things

IoT refers to uniquely identifiable objects and their virtual representations look similar to that of an Internet-like structure. IoT being a technological revolution represents the future of communication and computing. Connecting everything by IoT, processing and managing of massive data collected in the network for example in Big Bazaar Mall, where each item is tagged and extended to the whole world, could become possible for which an efficient network architecture is of prime importance. Similarly, in SIoT like Facebook, each member is connected to the other in the network in some way based on their taste and choice.

3.2 Trust and IoT

Let us take an example of traders association with user-specific information and policy construction. Traders' associations will control their group membership and users may choose to place default levels of trust in such organisations (the value 1.0 is chosen generally) and particularly familiar organisations. Groupings of friends could be established from social networking sites, phone books, etc. Similarly, consumer and interest groups would be found by the user through other activities and membership might cause a trust level question to be raised. Available groups from the web, e.g., search, keyword indexes, tags, maps of common interest, and suggesting groups based on aggregate data are considered with the following questionnaire, given in Figure 1. Generally, individuals express their trust through a percentage and less commonly with an absolute value. However, depending on the interactions nature of relations between individuals in a community, a meaningful way to represent the value of trust is properly designed.

Figure 1 Trade union's policy declaration example

if member of Kolkata Traders Association then trust = 1.0 (full trust)

if member of my Friends then trust = 1.0 (full trust)

if member of Ethical Consumers then trust = 0.8 (partial trust)

if max (rating (Ethical Consumers)) > 0.8 then trust = 0.5 (partial trust)

else trust = –1(dis-trust) or else

Trust = 0 (confused state of mind)

3.3 Trust models and trust computing

In this section, we provide an insight to the various trust models available followed by the computation techniques to make it a viable one. The trust representation model is shown in Figure 2. A similarity matching based on user and item rating to trust model is shown in Figure 3.

Figure 2 Trust representation

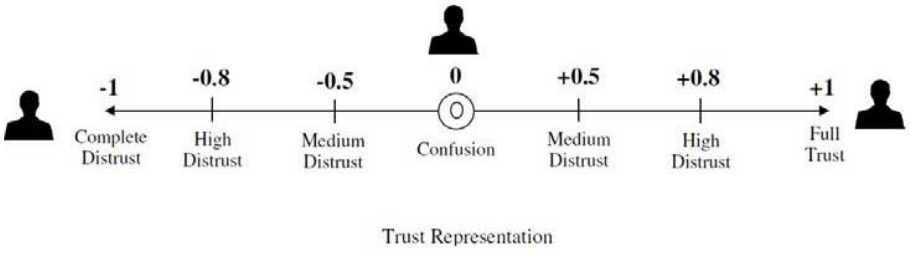
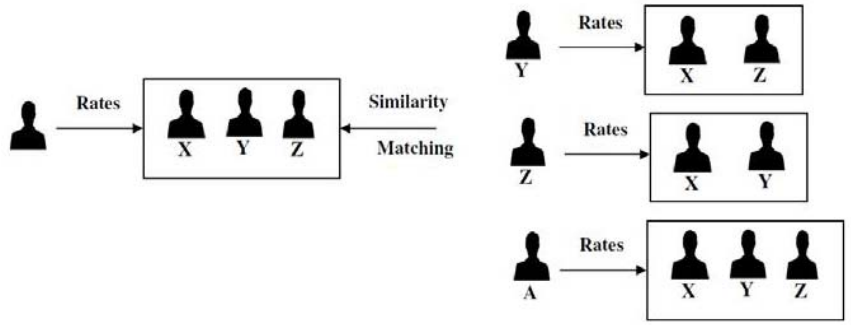
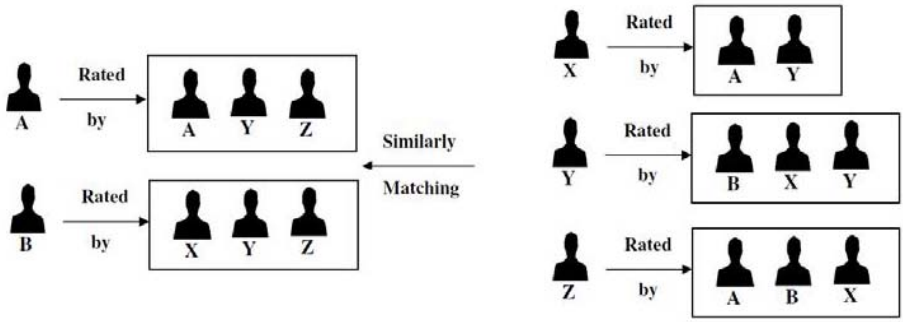


Figure 3 Similarity-based trust model, (a) user-based approach (b) item-based approach



(a)



(b)

3.3.1 Discrete trust models

Expressing trust in a scale of discrete data is easier to interpret in general as It would be simpler to say that a user is 'usually trusted' rather than expressing such statement as a percentage like trusted in 70% of cases (Jøsang, 2007). A community such as Epinion (epinions.com) uses a binary scale for the expression of trust: a user declares his trust in another (confidence) performed by the positive value of 1, or distrust by choosing the option of blocking that user and this would be interpreted as the negative value –1. The zero (represented on the graph of the network) indicates that there is no declared relationship yet trust between the two users. Computation of trust would eventually generate continues values (Guha et al., 2004); however, techniques for rounding the results within the followed scale are then introduced.

3.3.2 Probabilistic trust models

The main advantage from expressing trust with probabilities lies in developing models using advanced statistical methods, for their robustness in terms of treatment such as Bayesian approaches or models for reasoning through Markov chains (Patel et al., 2005). Thus, the probabilistic approaches can also be operated and the trust resulting value would be a continuous value proposed to the user helping for making his decision on whether trusting or not a different target user.

3.3.3 Belief models

In belief models, as proposed by Jøsang (2001); trust may be represented by a system holding a continuous value of trust, distrust and the uncertainty. Combining trust and distrust to represent the belief of a user at a given instant bring the model back to a belief model. This aspect appears in the propagation model of trust and distrust in Guha et al. (2004) where a belief matrix is set up combining all pairs of users' beliefs towards each other, a belief value that aggregates the portion of trust and distrust that a pair assigns to another.

3.3.4 Fuzzy models

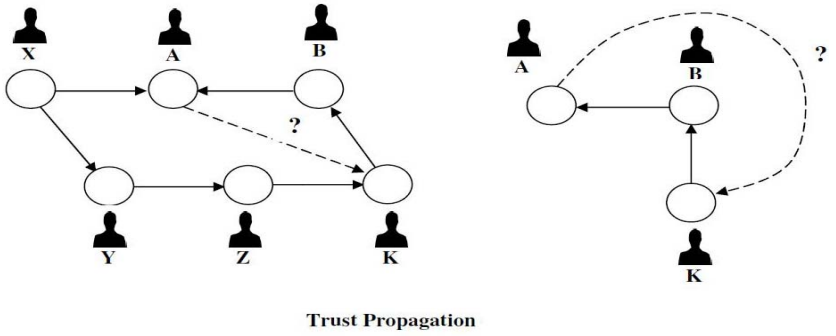
Fuzzy logic finds its suitability for trust evaluation as it takes the uncertainties in expressions used to determine the trust into account. Many authors (Aberer et al., 2006; Chen et al., 2005) propose a trust model for multi-agent system using fuzzy sets. Nefti et al. (2005) present a method based on fuzzy logic to evaluate trust in e-commerce arguing that fuzzy logic is suitable for trust evaluation as it takes into account the uncertainties within e-commerce data and like human relationships.

3.4 Computing trust

Trust value can be computed using two different approaches to distinguish between a pair of users. In the first case, trust computation is done considering propagation of values of trust through consecutive users' trust via an effective trust path leading to the target user, where as in the second case, this is obtained by analysing users' behaviours in a given social community (Abdessalem et al., 2010). Figure 4 shows the possible situations of potential transitions, which helps to propagate a value of trust between users as a part of a

web of trust with different possibilities of trust propagations (transitions) between users A and K.

Figure 4 Trust transitivity



Note: The dotted line indicates the trust to infer.

3.4.1 Distrust

A quantification and propagation approach to distrust is discussed first in Guha et al. (2004). The importance of distrust propagation arises by a third option in few trust-systems such as Epinion and eBay through which a user can manifest its distrust towards another by blocking him for example, beside classical options for expressing his trust in another (binary value 1). The zero value comes to reflect an initial state between any pair of users where no value of trust/distrust is yet revealed. In Quercia et al. (2007), the authors make the assumption that in this scale 0-1 a third value expressing the distrust can be integrated. However, reviewing the specification of the approach is necessary to ensure a generation of coherent values. In Guha et al. (2004), the authors discuss the possibility of depicting distrust through a negative value which might deteriorate the results as the expression of distrust is more informative than the trust in some cases.

4 Proposed methodology

4.1 Bayesian belief networks

BBNs (Pearl, 1988) are very effective for modelling situations where some information is already known and incoming data is uncertain or partially unavailable. Uncertainty arises in many situations. For example, experts may be uncertain about their own knowledge, there may be uncertainty inherent in the situation being modelled, or uncertainty about the accuracy and availability of information. Because BBN offer consistent semantics for representing uncertainty and an intuitive graphical representation of the interactions between various causes and effects, they are a very effective method of modelling uncertain situations that depend on cause and effect. Each of the variables in the BBN is represented by nodes. A variable in a belief network could be whether a light switch is on, the proximity of an enemy battalion, or the RPM of an engine. Each node has states, or a set of probable values for each variable. For example, the weather could be cloudy or

sunny, an enemy battalion could be near or far, symptoms present or not present, and the garbage disposal working or not working. Nodes are connected to show causality with an arrow indicating the direction of influence. These arrows are called edges. A BBN is a model that represents the possible states of a given domain. A BBN also contains probabilistic relationships among some of the states of the domain. The probability of any node in the BBN being in one state or another without current evidence is described using a conditional probability table. Probabilities on some nodes are affected by the state of other nodes, depending on causality. Prior information about the relationships among nodes may indicate that the likelihood that a node is in one state is dependent on another node's state. With the historical information stored in the conditional probability tables, BBN can be used to help make decisions, or as a way of automating a decision-making process.

4.2 Fuzzy expert system

A fuzzy expert system (Zadeh, 2002) is an expert system that uses a collection of fuzzy membership functions and rules, instead of Boolean logic, to reason about data. The rules in a fuzzy expert system are usually of a form similar to the following:

If A is low and B is high then O = medium where A and B are input variables, O is an output variable, low is a membership function (fuzzy subset) defined on A, high is a membership function defined on B, and medium is a membership function defined on O. The antecedent describes to what degree the rule applies, while the conclusion assigns a membership function to each of one or more output variables. Most tools for working with fuzzy expert systems allow more than one conclusion per rule. The set of rules in a fuzzy expert system is known as the rule base or knowledge base. The general inference process proceeds in following steps.

1 Under *fuzzification*, the membership functions defined on the input variables are applied to their actual values, to determine the degree of truth for each rule antecedent.

2 Under *inference*, the truth value for the premise of each rule is computed, and applied to the consequent part of each rule. This results in one fuzzy subset to be assigned to each output variable for each rule.

3 Under *composition*, all of the fuzzy subsets assigned to each output variable are combined together to form a single fuzzy subset for each output variable.

4 Finally is the (optional) *defuzzification*, which is used when it is useful to convert the fuzzy output set to a crisp number.

4.3 Conditional independence

It is often stated that tackling the task of selecting a Bayesian network structure from data consists of two distinct approaches (Cowel et al., 1999): firstly, apply CI tests. When testing for the presence or otherwise of edges; and secondly search the model space using a scoring metric. CI is a highly important concept in statistics and artificial intelligence. Properties of probabilistic CI provide theoretical justification for the method of local computation (Cowel et al., 1999) which is at the core of probabilistic expert systems (Jensen, 2001), successfully applied in numerous areas. The importance of CI is

given by its interpretation in terms of relevance among symptoms or variables in consideration (Pearl, 1988; Studeny, 2011); that are why it is crucial in probabilistic reasoning.

4.3.1 Advantages

The key advantage of assuming that features are class-conditionally independent is that it reduces the curse of dimensionality (Jarecki et al., 2013). For example, for ten binary features there are 2^{10} possible feature configurations and a need to estimate 1,024 likelihoods of feature configurations for each class. Assuming class-CI reduces the number of required likelihoods from 1,024 to 8. Another benefit is that class-CI allows inferences about new feature configurations. Even if a particular combination of feature values has not been observed yet, assuming class-CI allows inference of the likelihood of the feature configuration from the marginal likelihoods of the individual feature values, thereby enabling computing the posterior class probabilities.

4.3.2 Robustness

While class-CI may rarely exactly hold in real-world environments, violations of this assumption do not necessarily impair performance. For instance, a widely used classifier is the naive Bayes model, which treats features as class-conditionally independent and computes the posterior class probabilities accordingly. Both simulation studies and analytic results demonstrate the robustness of this model under a variety of conditions (Jarecki et al., 2013). For instance, if the optimality criterion is classification accuracy, then even if the derived posterior probabilities do not exactly correspond to the true posterior, as long as the correct category receives the highest posterior probability, classification error will be minimised.

4.4 K-nearest neighbour

K-nearest neighbour (K-NN) classifier is one of the simplest classifier that discovers the unidentified data point using the previously known data points (nearest neighbour) and classified data points according to the voting system (Silver et al., 2001). K-NN classifies the data points using more than one nearest neighbour. K-NN has a number of applications in different areas such as health datasets, image field, cluster analysis, pattern recognition; online marketing, etc.

K-NN (Arroyo and Mate, 2009) is a non-parametric instance-based learning as it allows a hypothesis of model complexity to grow with data sizes. K-NN is based on minimum distance from a query instance to all training samples to determine the K-NN, which span the entire input space. The Euclidean distance of lower dimensional space is commonly applied for computing the minimum distance in this step. Prediction of the query instance is taken as majority votes of the K-NNs. The idea is that any point A is likely to be similar to those points in the neighbourhood of A. The choice of parameter value K is critical but K-NN is advantageously robust to uncertainty or noisy training samples.

A decision rule for KNN classifier is very simple and can be generalised for any number of classes (Khenchaf and Hoeltzener, 2012). It is a non-parametric probabilistic approach. In this method, the only parameters to be determined are the parameter k and

the distance measure used to compare the subject to recognise and find the nearest neighbouring objects.

4.5 Fuzzy K-NN

While the fuzzy K-nearest neighbour (fuzzy K-NN) procedure is also a classification algorithm the form of its result s differs from the crisp version. The fuzzy K-NN algorithm assigns class membership to a sample vector rather than assigning the vector to a particular class. The advantage is that no arbitrary assignments are made by the algorithm. For example, if a vector is assigned 0.9 memberships in one class and 0.05 memberships in two other classes we can be reasonably sure the class of 0.9 memberships is the class to which vector belongs. On the other hand, if a vector is assigned 0.55 memberships in class one, 0.44 memberships in class two, and 0.01 membership in class three, then we should be hesitant to assign the vector based on these results. However, we can feel confident that it does not belong class three. In such a case the vector might be examined further to determine its classification, because the vector exhibits a high degree of membership in both classes one and two. Clearly the membership assignments produced by the algorithm can be useful in the classification process (Keller et al., 1985).

4.5.1 Fuzzy K-NN algorithm

The pseudo code of the fuzzy K-NN with K = 10 is shown in Figure 5.

Figure 5 Pseudo code of fuzzy K-NN algorithm

Inputs	Input patterns of training set (PTR), labels of training set (T), number of neighbours (k = 10), patters of testing set (PTE)
Output	classification vector of testing set (y)
Step 1	Initialisation: Set i = 0, m = z (let, z = 2, as default fuzzy parameter)
Step 2	N ← number of data elements in PTR
Step 3	n ← number of data elements in PTE
Step 4	q ← number of unique values of T (classes)
Step 5	transform T into a q × N binary matrix (consisting of a row for each pattern, in which there is a 1 in the column corresponding to the class it belongs to, and 0 in all other places)
Step 6	do i ← (i + 1)
Step 7	calculate distance array (d) based on the Euclidean distances between test pattern i and each one of the training patterns
Step 8	sort distances and store indexes
Step 9	get k nearest patterns based on the first k values of array
Step 10	set weights array (w) equal to membership function
Step 11	if a value of w is infinite then replace this value with 1
Step 12	calculate memberships: M ← T * wT / Σw for all the neighbours (based on index array)
Step 13	Q ← j value for which M_j is maximised
Step 14	classify pattern i to class Q (y_i ← class Q)
Step 15	until i = n

Here, $\mu_{ij}(y)$ can be assigned class membership in several ways. They can be given complete membership in their known class and non-membership in all other. Here, this classifier makes use of an objective function of the distance of each pattern to each neighbour. These functions take the form of weights which are calculated as $w(i) = \frac{1}{d_{ij}}$ and distance d measured from pattern (i) to its neighbour (j). The weights are then normalised and finally, combined with the class labels of the neighbours to obtain the classification output.

5 Experimental data

We use publicly available social network trust dataset such as Epinion dataset and Slashdot Zoo network dataset for our proposed model.

5.1 The Epinion dataset

This is one of the most popular datasets for scientific communities interested on trust computing. TrustLet (Massa and Avesani, 2006), for example, is a cooperative environment for the scientific research of trust metrics on social networks. It gives the opportunity to researchers to compare all proposed trust metrics on the same datasets. As our approach aims to reach best performances then achievements, we worked on the same Epinion Dataset given for free download from the TrustLet website. Epinion is a website where people can review products. Users can register for free and start writing subjective reviews about many different types of items (music, hardware, software, office appliances, television show, etc.). A peculiar characteristic of Epinion is that users are paid according to how much a review is found useful (income share programme).

In this environment, the attempts to game the systems are many and, as a possible fix, a trust system was put in place. Users can add other users to their 'web of trust', i.e., "reviewers whose reviews and ratings they have consistently found to be valuable", and their 'block list', i.e., "authors whose reviews they find consistently offensive, inaccurate, or in general not valuable". Our experimental data sample was picked from Epinion's web of trust files given at TrustLet and described in the following.

- *User_rating* file: Trust is the mechanism by which the user makes a statement that he likes the content or the behaviour of a particular user, and would like to see more of what he/she does in the site. Distrust is the opposite of the trust, in which the user says that he/she do want to see lesser of the operations performed by a particular user. The column details:
  1 MY_ID, this stores Id of the member who is making the trust/distrust statement
  2 OTHER_ID, the other ID is the ID of the member being trusted/distrusted
  3 VALUE, equal to 1 for trust and –1 for distrust
  4 CREATION, it is the date on which the trust was made explicitly.
- *mc* file: contains information on each article written by a user. The column details include:
  1 CONTENT_ID is an object ID for the article

2 AUTHOR_ID is the ID of the user who wrote the article

3 SUBJECT_ID is the ID of the subject that the article is supposed to be about.

- *rating* file: Ratings are quantified statements made by users regarding the quality of a content in the site. Ratings are the basis on which the contents are sorted and filtered. The column details:

  1 OBJECT_ID is the ID of the object being rated. The only valid objects considered up to now are the reviews and essays (identified their content_idin a member_content table).

  2 MEMBER_ID stores the id of the member (user) who is rating the object

  3 RATING stores the 1–5 rating (1 – not helpful, 2 – somewhat helpful, 3 – helpful, 4 – very helpful, 5 – most helpful) of the considered object by a given member

  4 STATUS antiquates the display status of the rating: 1 means the member has chosen not to show his rating of the object, and 0 means that the member does not mind showing his name besides the rating.

  5 CREATION indicates the date on which the member first rated the object

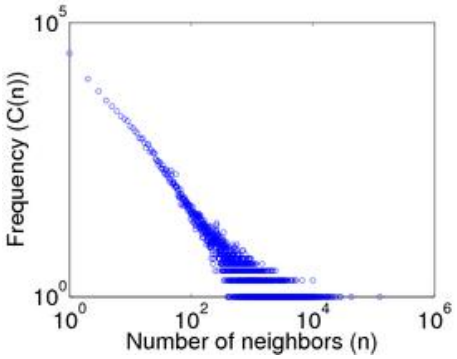
  6 LAST_MODIFIED is the latest date on which the member modified his rating of the object

  7 TYPE is not used up to now. When Epinion will allow more than just content rating to be stored in this table, then this column would store the type of the object being rated.

  8 VERTICAL_ID of the review.

A degree distribution of the Epinion data with user rating is shown in Figure 6.

Figure 6 Degree distribution of user rating in Epinion dataset (see online version for colours)

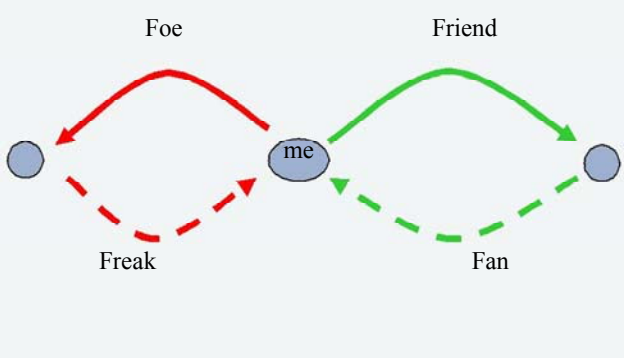


The trust files can be viewed as a directed graph. The data they contain consists of 131 829 nodes and 841 372 edges, each labelled either trust or distrust. Of these labelled edges, almost 85% are labelled trust. We interpret trust to be the value +1 and distrust to be –1.

5.2 Slashdot Zoo network dataset

This is the signed social network of users of the technology news site Slashdot (slashdot.org) (Kunegis, 2009), connected by directed 'friend' and 'foe' relations, as shown in Figure 7. The 'friend' and 'foe' labels are used on Slashdot to mark users, and influence the scores as seen by each user. For instance, If user A marks user B as a foe, the score of user B's posts will be decreased as shown to user A.

Figure 7 A scenario of Slashdot Zoo dataset (see online version for colours)



Slashdot is a technology news platform where users can post and read other users' news articles and comments. On Slashdot, users can create two types of explicit and directed social links between themselves and other users. These are labelled friend and foe. Both link types allow the user to change the visibility of the content the linked user has created. Although the effect of a link is not predetermined but user configurable the convention is that the friend link increases the content visibility, the foe link decreases content visibility of the target user. Therefore the friend link is a positive link, while the foe link is a negative link. The friend and foe link types are also called fan and freak from the point of view of the targeted user.

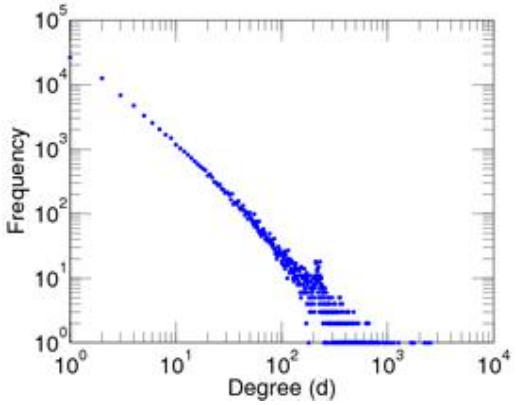
In some online social media such as Slashdot, actors are allowed to explicitly show their trust or distrust towards each other. Such a network, called a signed network, contains positive and negative edges. Traditional notions of assortativity and disassortativity are not sufficient to study the mixing patterns of connections between actors in a signed network, owing to the presence of negative edges (Rathore et al., 2013).

We use two large online social networks Epinion and Slashdot where each link is explicitly labelled as positive or negative. All these networks are downloaded from Stanford large network dataset collection. Since the original graphs are too large and sparse, we select 20,000 nodes from Epinion and 16,000 nodes from Slashdot with the highest degrees, as well as the edges between the selected nodes. There are 13 nodes in Epinion and 1 node in Slashdot that are disconnected from the remaining selected nodes. These isolated nodes are removed from the respective network and the remaining ones form a connected component. Figure 8 shows the degree distribution of the Slashdot Zoo dataset. Table 1 shows the statistics of the extracted networks.

Table 1 Statistics of extracted graphs

Parameters	*Epinion*	*Slashdot*
Number of nodes	19,987	15,999
Number of edges	634,215 371	122 121
Average degree	31.731	23.197
Positive edges	87.6%	76.5%
Average distance	3.163	3.569

Figure 8 Degree distribution in Slashdot network dataset (see online version for colours)



We can observe that Epinion has the largest number of nodes, edges, average degree and the percentage of positive edges than the Slashdot networks. The statistics demonstrate that there indeed exists discrepancy in data distribution in various networks.

Table 1 summarises the two datasets that have both positive and negative links between users, forming a directed, asymmetric signed network.

Although the functionality that lies behind the link types is not fully identical between Slashdot and Epinion, it is very similar according to our definition of positive and negative links. Hence, based on this similar functionality we assume similar properties of the two networks motivated us to use both these datasets for experiments in building our proposed trust management model.

Based on these premises, this paper investigates various data mining problem of learning negative from positive links in a signed network. This problem is related to the link prediction consisting of predicting future edges in an un-weighted network. In the two social networks Slashdot and Epinion with negative links, we study to what accuracy negative links can be inferred by positive links.

5.3 Trust metrics used in the experiments

The following are some of the trust metrics used for evaluation of the proposed trust management model.

5.3.1 Accuracy

The accuracy of classifiers is the percentage of correctness of prediction among the test sets. This refers to the ability of the model to correctly predict the class label of new or previously unseen data.

5.3.2 Mean absolute error

Calculation of MAE is relatively simple. It involves summing the magnitudes (absolute values) of the errors to obtain the 'total error' and then dividing the total error by n; once again, assuming that the wi s are all equal to 1.0.

5.3.3 Root mean square error

Root mean square error (RMSE) is computed with three simple steps, where at first 'total square error' is obtained as the sum of the individual squared errors; secondly, total square error then is divided by *n*, which yields the mean square error (MSE). The third and final step is to take RMSE as the square root of the MSE.

5.3.4 Coverage

Coverage simply refers to the fraction of ratings for which, after being hidden, the proposed algorithm is able to produce a predicted rating. It might in fact for some cases the proposed methodologies are not able to predict the user rating would give the rating to an item. In this paper, we use users' coverage, which is defined as the portion of users for which the algorithm is able to predict at least one rating. Actually, we focus on model for its suitability in predicting all the ratings for a user who provides many ratings and performs poorly for a user who has rated few items.

5.3.5 F-measure

F-measure is used as one of the performance measure in the proposed approach. While precision and recall are a standard measure for exactness and completeness, respectively, F-measure is the harmonic mean of precision and recall.

5.3.6 Trust and similarity metric

In order to measure similarities, the proposed model reports trustee and trusters in terms of nodes and their relationship in terms of edges. The criteria here is trusters A and B are similar only when Trustees C and D themselves are similar, then A and B are somewhat similar to C and D. Similarity matrix is defined as $S_{ij} = 1$ when the actors i, j know each other, and $S_{ij} = 0$ otherwise. Trust matrix T describes reciprocal trust of actors. Existing trust is given by value in the interval (0, 1). Value –1 represents the situation when the actors do not know each other or the fact that reciprocal trust is not known.

5.3.7 Degree, betweenness, closeness

Degree, closeness, betweenness and Eigen vector are referred as centrality measures for individual nodes during social network analysis.

- *Degree:* It defines the number of ties a node has (also called nominations) in a directed network: in-degree and out-degree In-degree (number incoming ties) also called prestige.
- *Closeness:* It is based on the average distance from a node to every other reachable node in the network. It is calculated as the inverted sum of the shortest paths between the node and every other node.
- *Betweenness:* This depends on the number of cases in which a node lies on the shortest path between two other nodes in the network that is adjusted by total number of shortest paths. As quite evident that while degree and closeness centrality depicts the reachability of a person, betweenness centrality gives an idea how a person is more important if he/she was more intermediary in the network. The more a person is a go-between, the more central his/her position in that network. This reflects the importance of a person being in the middle of social communications of a network and to what extent he/she is needed as a link in the chains of contact in the society. On the other hand, a vertex has betweenness centrality = 0 if it was not located between any other vertices in the network, which points out to a weak social role that he/she plays.
- *Eigenvector centrality:* This centrality of the node can be determined by taking the centralities of the nodes to which the node is adjacent. In order to normalise, the score is divided by the square root of one half.

5.3.8 Trust transitivity

Trust transitivity means, for example, that if A trust B whose trust K, then A will also trust K, with an assumption that A is aware of the trust relation between B and K, which may not be the case in real life always. However, a path A-B-K of length two from a graph as shown in Figure 5 is transitive if A is connected to K. An unordered triple is transitive if it contains a transitive path, where transitivity is the number of paths (triples) which are transitive divided by the number of triples which have the potential to be transitive by the addition of a single edge. In this experiment, we choose the software to detect automatically whether the data is directed or undirected with standard transitivity method to obtain the following:

a Base: it represents the number of potentially transitive paths or triples. That is either number of paths of length two or the number of triples containing paths of length two. Here, we use the later one for our experiments.

b Open: this gives the number of base paths or triples that were not transitive.

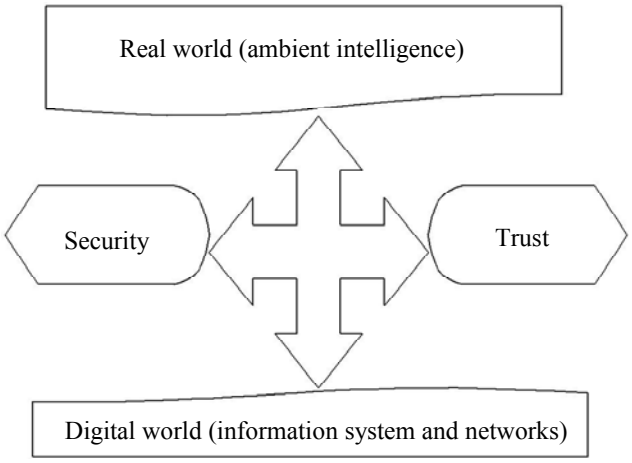
c Closed: This is the number of base paths or triples that are transitive It is noted that Base = Open + closed.

However, in some cases, trust can be transitive and can be used to derive trust (Josang and Pope, 2005; Christiansen and Harbison, 2003).

6 Experiments and results

A general framework for developing a SIoT Trust model is shown in Figure 9.

Figure 9 A general framework for SIoT trust model



The real world characterises the social relationships such as: friendship, ownership and community, who are interested to communicate each other through the digital world consisting of information systems and network. While maintain relationships in this way, one need to trust the other and the care should be taken that the information shared among them is secured. This very important, as uncooperative node will act as they wish and the malicious nodes with try to use the intrusive behaviour to perform the unlawful activities. Hence, developing a reliable trust management model using data mining techniques became inevitable for the SIoT, which is our main focus in this paper.

While building trust model for the IoT, we use trust in the range [–1; +1], the value +1 indicating complete (blind) trust1, the value –1 indicating complete distrust, and the value 0 indicating indecision (i.e., more information is needed).

The complexity in measuring trust score and predicting trustworthiness in social IoT networks is most promising and leads to many problems. These include how to quantify the capability of individual devices in the trust dynamics and how to assign concrete level of trust in user to user communication. Also trust relationship in IoT environment is hard to ascertain due to uncertainties involved. The benefits of fuzzy trust calculations includes: This inferences using fuzzy approach can easily quantify uncertainties for the measuring the level of trust in uncertain IoT environment. Further, it is easy to develop membership function and inference rules for different trust relationship using fuzzy approach. Another advantage of fuzzy approach as compare to the other approaches is that it can handle incomplete and imprecise inputs in decentralised environment where resource owners usually do not have complete and precise inputs. Finally, Fuzzy approach is flexible, intuitive knowledge-based tool which is easy for computation and validation.

Fuzzy k-NN algorithm assigns class memberships to a sample vector rather than assigning the vector to a particular class, with an obvious advantage of not making any arbitrary assignments where the resultant classification is assured by vector membership values.

We use Fuzzy fusion of K-NN and CI search in Bayesian belief propagation and the BBN with CI to design the proposed trust model. All the experiments are conducted in an Intel Pentium PC with 2.6 GHz CPU, 2 GB RAM and 500 GB HDD. We use Java-based data mining tool (Witten and Frank, 2005) with ten-fold cross validation for the proposed work.

6.1 Results and discussion

The visualisation of Slashdot and Epinion networks used in this paper are shown in Figures 10 and 11. Social network consists of many actors with their trust relations. Actors and their known contacts in social network are given by the similarity matrix of actors S. Matrix entries S_{ij}, S_{ji} = 1 when the actors i, j know each other, and S_{ij} = 0 otherwise. This is shown for sample Epinion and Slashdot Zoo datasets in Tables 2 and 3 respectively. Trust matrix T describes reciprocal trust of actors. Existing trust is given by value in the interval (0, 1). Value –1 represents the situation when the actors do not know each other or the fact that reciprocal trust is not known. The sample trust matrix for Epinion and Slashdot Zoo dataset are shown in Tables 4 and 5 respectively.

Figure 10 Visualisation of Slashdot network (see online version for colours)

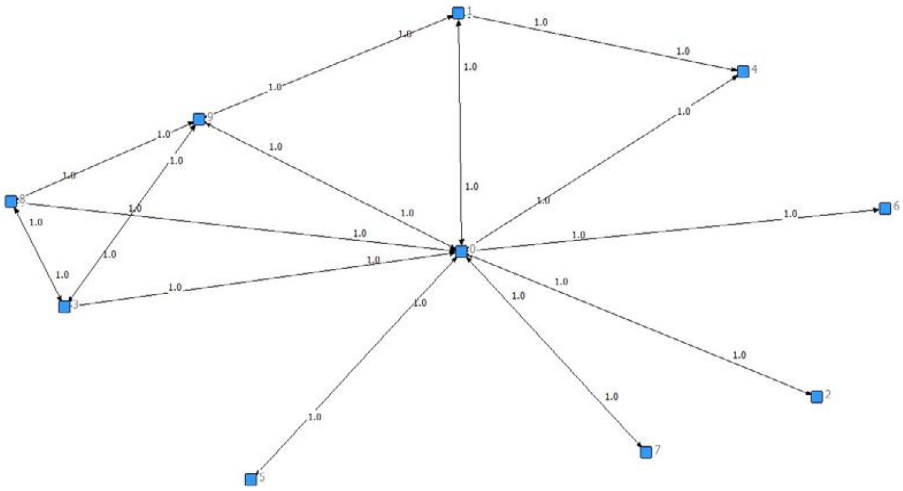


Figure 11 Visualisation of Epinion dataset (see online version for colours)

Table 2 Epinion dataset similarity based on correlation

From/to	*2*	*17*	*15*	*11*	*1,109*
Node					
2	1.000	0.000	0.000	0.000	0.000
17	0.000	1.000	1.000	–0.500	–0.500
15	0.000	1.000	1.000	–0.500	–0.500
11	0.000	–0.500	–0.500	1.000	1.000
1,109	0.000	–0.500	–0.500	1.000	1.000

Table 3 Similarity matrix based on correlation for Slashdot Zoo dataset

From/to	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Node										
0	1.000									
1		1.000	0.509	.524	0.218	0.509	0.509	0.509	0.524	-0.089
2		0.509	1.000	0.509	0.667	1.000	1.000	1.000	0.509	0.408
3		0.524	.509	1.000	0.218	0.509	0.509	0.509	0.524	0.356
4		0.218	.667	0.218	1.000	0.667	0.667	0.667	0.218	0.612
5		0.509	1.000	0.509	0.667	1.000	1.000	1.000	0.509	0.408
6		0.509	1.000	0.509	0.667	1.000	1.000	1.000	0.509	0.408
7		0.509	1.000	0.509	0.667	1.000	1.000	1.000	0.509	0.408

Table 3 Similarity matrix based on correlation for Slashdot Zoo dataset (continued)

From/to	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Node										
8		0.524	0.509	0.524	0.218	0.509	.509	0.509	1.000	0.356
9		–0.089	0.408	0.356	0.612	0.408	0.408	0.408	0.356	1.000

Table 4 Trust matrix for Epinion dataset

	2	*17*	*15*	*11*	*1,109*
1	–1	0	0	0	0
7	0	0	0	0	0
16	0	1	0	0	0
26	0	0	–1	1	1
99	0	0	0	0	0

Table 5 Trust matrix for Slashdot

	0	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
0	1	1	1	1	1	1	1	1	1	1
1	1	0	0	0	1	0	0	0	0	1
2	0	0	0	0	0	0	0	0	0	0
3	1	0	0	0	0	0	0	0	1	1
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
6	1	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0
8	1	0	0	1	0	0	0	0	0	1
9	0	0	0	0	0	0	0	0	0	0

As expected, the basic user/node similarity matrix S has the following properties:

- Due to the positivity property, all values of S are positive numbers in the interval [0, 1]. The maximum value of similarity (equal to 1) arises, when two nodes are connected with only one edge and have no connections with other nodes, whereas the minimum value (equal to 0) arises when two nodes do not share any edge.
- Due to the reflexivity property, all values of the main diagonal in S are equal to 1.
- Due to the symmetry property, S is a symmetric square matrix.

In order to analyse the social network model, we use two-mode centrality measures for both the datasets, which are outlined in Tabled 6 and 7 for Epinion and Tables 8 and 9 for Slashdot Zoo network.

Table 6 Two-mode centrality measures for *rows* of Epinion

	Degree	*Closeness*	*Betweenness*	*Eigenvector*
1	0.000	0.000	0.000	0.000
3	0.100	0.609	0.000	0.102
5	0.000	0.000	0.000	0.000
7	1.000	1.400	0.617	0.974
10	0.100	0.609	0.000	0.102
12	0.100	0.609	0.000	0.102
14	0.100	0.609	0.000	0.102
16	0.100	0.609	0.000	0.102
18	0.000	0.000	0.000	0.000
20	0.000	0.000	0.000	0.000

Table 7 Two-mode centrality measures for *columns* of Epinion

	Degree	*Closeness*	*Betweenness*	*Eigenvector*
2	0.100	0.824	0.000	0.300
4	0.200	0.875	0.086	0.332
6	0.100	0.824	0.000	0.300
8	0.100	0.824	0.000	0.300
9	0.100	0.824	0.000	0.300
11	0.200	0.875	0.086	0.332
13	0.200	0.875	0.086	0.332
15	0.200	0.875	0.086	0.332
17	0.200	0.875	0.086	0.332
22	0.100	0.824	0.000	0.300

Table 8 Two-mode centrality measures for ROWS of Slashdot Zoo

	Degree	*Closeness*	*Betweenness*	*Eigenvector*
0	1.000	1.000	0.553	0.656
1	0.300	0.667	0.036	0.294
2	0.100	0.609	0.000	0.154
3	0.300	0.667	0.029	0.316
4	0.200	0.636	0.014	0.223
5	0.100	0.609	0.000	0.154
6	0.100	0.609	0.000	0.154
7	0.100	0.609	0.000	0.154
8	0.300	0.667	0.029	0.316
9	0.400	0.700	0.056	0.372

Table 9 Two-mode centrality measures for COLUMNS of Slashdot Zoo

	Degree	*Closeness*	*Betweenness*	*Eigenvector*
0	1.000	1.000	0.553	0.656
1	0.300	0.667	0.036	0.294
2	0.100	0.609	0.000	0.154
3	0.300	0.667	0.029	0.316
4	0.200	0.636	0.014	0.223
5	0.100	0.609	0.000	0.154
6	0.100	0.609	0.000	0.154
7	0.100	0.609	0.000	0.154
8	0.300	0.667	0.029	0.316
9	0.400	0.700	0.056	0.372

From Tables 6 to 9, the following observations are made:

- We can see that we have nodes with degree centrality = 0 because these are 'isolates'. The larger the value, the more central a user in the network in regard to degree centrality. It can be seen that node 7 in Epinion and Node 0 in Slashdot Zoo datasets are more central than the others.
- Closeness centrality values range from 0 (for isolated vertices) to
  1 It is easy to notice that node 0 in Epinion and node 7 in Slashdot Zoo dataset are the top closeness centrality user.
- It is easy to conclude that those nodes have betweenness centrality = 0 were not located between any other nodes in the network, which points out to a weak social role that the user plays. The node 7 in Epinion with a value of 0.617 and 0.553 for node 0 in Slashdot Zoo has more social role in comparison to the others.
- Further, eigenvector centrality, presents how well one user is connected to other as a result of their structure of shared social relations, would be positively related. Here, once again the node 7 and node 0 for Epinion and Slashdot Zoo dataset steal the show with highest value among others.

It can be very well noticed that all the four measures (degree, closeness, betweenness and Eigen vector centrality) have showed similar (not identical) results, which support the notion that all these measures collectively are used to measure most important individuals in a community relationships.

As discussed earlier, we also perform trust transitivity analysis to obtain a better trust management model, the result obtained thereto are provided in Table 10.

Table 10 Transitivity for Epinion dataset and Slashdot Zoo dataset

Standard transitivity	*Epinion*	*Slashdot Zoo*
Statistics/dataset		
Base	0.000	40.000
Open	0.000	34.000
Closed	0.000	6.000
Transitivity	nil	0.150

Table 11 Evolutionary algorithms on trust model evaluations

Dataset/algorithm	*Slashdot Zoo network dataset*					*Epinion dataset*				
	Acc	*MAE*	*F-measure*	*Coverage*	*RMSE*	*Acc*	*MAE*	*F-measure*	*Coverage*	*RMSE*
Tidal trust (Jamali and Ester, 2011)	-	-	-	-	-	-	-	0.77	82.36	1.109
Mole trust (Jamali and Ester, 2011)	-	-	-	-	-	-	-	0.765	81.03	1.104
Trust walker (Jamali and Ester, 2011)	-	-	-	-	-	-	-	0.819	93.22	1.079
DiffTrust (Fang et al., 2013)	-	-	-	-	-	0.8163	0.206	-	-	-
Baseline (Fang et al., 2013)	-	-	-	-	-	0.6365	0.342	-	-	-
NB+Random guess (Quercia et al., 2007)	-	-	-	-	-	0.311	-	-	-	-
Asso. rules with 10% supp and 90% conf (Bachi et al., 2012)	0.22	-	-	-	-	0.97	-	-	-	-
Ours, BBN+CI	0.87	0.181	0.923	98.99	0.3061	0.898	0.156	0.936	98.07	0.2824
Ours, fuzzy K-NN+BBN	0.83	0.161	0.905	83.79	0.4026	0.8451	0.232	0.898	97.95	0.3424

From Table 10, we can observe that overall transitivity is 0 for Epinion and 0.150 for Slashdot Zoo dataset, concluding that networks with high level of transitivity are often more stable, balanced, harmonious.

Finally, we perform data mining approach to obtain a reliable trust management model and the obtained result for the Epinion and Slashdot Zoo network datasets along with other's obtained results are summarised in Table 11.

In diffTrust model, the authors adopt other buyers direct trust evaluations on an advisor derived from their shared interactions with the advisor by considering these buyers social proximity with respect to the current buyer, which assures more accurate trust evaluation of the advisor, in comparison to the baseline approach.

Ours is better than (Quercia et al., 2007; Fang et al., 2013), as the later ones could not able to disclose their trust in other for both centralised and decentralised communities. In comparison to association rule methods used (Bachi et al., 2012), we achieve high accuracy in case of Slashdot Zoo dataset but less in Epinion dataset. The reason behind is that the former uses a high confidence threshold values which does not consider many association rules, hence many edges do not have a candidate rule, so they are not classified. From this, we can conclude that this method may not be a good choice.

In our approach, we receive different trust evaluations based on negative ratings, equally treating the positive ratings obtained from other buyer's direct trust evaluations towards the advisor.

From all the comparisons, it seems that our proposed approach outperforms other approaches, consistently achieving high accuracy, low MAE, low RMSE, high coverage, and high F-score demonstrating the effectiveness of our proposed trust management model.

The BBN is implemented with a motive to have dimensionality reduction in feature space, so that performance of the classifier (here, we use fuzzy k-NN) can be improved by discarding the 'noise' features.

It is observed that there is degradation in the accuracy, as from the point of view of BBN classifier alone, through the addition of superfluous features, but still found best in comparison to other existing approaches. In view of trivial algorithms using all the above, in order to make sure of ourselves for the effectiveness of its implementation, we further use statistical significance test of fuzzy BBN with Bayesian network with CI and association rule algorithms in terms of accuracy. We perform non-parametric Friedman two tail test with significance level alpha = 0.05. As the computed p-value = 0.607 is more than the alpha value, we cannot reject the null hypothesis. Hence, we conclude that both are significant in developing the trust model, even with difference in predictive accuracy.

7 Conclusions

The management of trust as well as access control in a web of trust are becoming key issues for many social websites. The idea of estimating a direct trust rate between two actors in a social network is probably a fast and effective way. However, robust and accurate inference techniques for the calculation of such measures are necessary, given the number of constraints that could affect the accuracy of the result. This paper gives an overview of the existing work and approaches to the problem of trust inference and lists the methods followed for this aim. We try to describe the experiments on the Epinion and

Slashdot Zoo datasets using hybrid fuzzy K-NN+BBN and BBN alone with various performance measures such as coverage and Transitivity apart from accuracy, MAE, RMSE and F-measure. Further, we use two tailed statistical significance test for understanding the importance of the used classifiers. These experiments enabled us evaluating the most effective methods of trust inference proposed in the literature and to test new extensions and refinements of existing approaches.

References

Abdessalem, T., Cautis, B. and Souhli, A. (2010) *Trust Management in Social Networks*, ISICIL Project ANR-08-CORD-011-05.

Aberer, K., Despotovic,Z., Galuba, W. and Kellerer, W. (2006) 'The complex facets of reputation and trust', in *9th Fuzzy Days, International Conference on Computational Intelligence Fuzzy Logic Neural Networks Evolutionary Algorithms*, pp.283–294.

Al-Masri, E. and Mahmoud Q.H. (2009) 'Discovering the best web service: a neural network-based solution', in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, pp.4250–4255.

Arroyo, J. and Mate, C. (2009) 'Forecasting histogram time series with k-nearest neighbour methods', *International Journal of Forecasting*, Vol. 25, No. 1, pp.192–207.

Avesani, P., Massa, P. and Tiella, R. (2005) 'Moleskiing.it: a trust-aware recommender system for ski mountaineering', *International Journal for Infonomics*, Vol. 20, No. 35, pp.1–10.

Bachi, G., Coscia, M., Monreale, A. and Giannotti, F. (2012) 'Classifying trust/distrust relation in online social network', in *Proc. of International Conference on Social Computing (Social Com)*, pp.552–557, IEEE.

Caverlee, J., Liu, L. and Webb, S. (2008) 'Socialtrust: tamper-resilient trust establishment in online communities', in *JCDL*, pp.104–114.

Chen, G., Li, Z., Cheng, Z., Zhao, Z. and Yan, H. (2005) 'A fuzzy trust model for multi-agent system', in *ICNC*, Vol. 3, pp.444–448.

Christiansen, B. and Harbison W.S. (2003) 'Why is not trust transitive?', in *Proc. of the 6th International Conference on Information Fusion*.

Coetzee, L. and Eksteen, J. (2011) 'The internet of things – promise for the future? An introduction', in Paul Cunningham and Miriam Cunningham (Eds.): *IIMC International Information Management Corporation*, pp.1–9, ISBN: 978-1-905824-24-3.

Cowell, R.G. (2001) 'Conditions under which conditional independence and scoring methods lead to identical selection of bayesian network models', *UAI2001*, pp.91–97.

Das, A. and Islam, M.M. (2012) 'SecuredTrust: a dynamic trust computation model for secured communication in multi-agent systems', *IEEE Trans. Dependable and Secure Computing*, Vol. 2, No. 2, pp.261–274.

Dasgupta, P. (1990) 'Trust as a commodity', in D. Gambetta (Ed.): *Trust: Making and Breaking Cooperative Relations*, Basil Blackwell, Oxford.

Deutsch, M. (2004) *Distributive Justice: A Social Psychological Perspective*, Yale University Press, USA.

DuBois, T., Golbeck, J. and Srinivasan, A. (2009) 'Rigorous probabilistic trust inference with applications to clustering', in *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, IEEE Computer Society, Vol. 1, pp.655–658.

Fang, H., Zhang, J. and Thalmann, N.M. (2013) A trust model stemmed from the diffusion theory for opinion evaluation', in Ito, Jonker, Gini, and Shehory (Eds.): *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013)*, May 6–10, Saint Paul, Minnesota, USA.

Golbeck, J. (2006) 'Combining provenance with trust in social networks for semantic web content filtering', *Proceedings of the International Provenance and Annotation Workshop*, Chicago, Illinois, USA, Vol. 6, pp.101–108.

Golbeck, J. (2006) Trust on the World Wide Web: A survey. Foundations and Trends in Web Science, 1(2):131–197.

Goyal, A., Bonchi, F. and Lakshmanan, L.V. (2010) 'Learning influence probabilities in social networks', in *WSDM '10: Proceedings of the Third ACM International Conference on Web Search and Data Mining*, ACM, NewYork, NY, USA, pp.241–250.

Guha, R., Ravi Kumar, Raghavan, P. and Tomkins, A. (2004) 'Propagation of trust and distrust', in *WWW '04: Proceedings of the 13th International Conference on World Wide Web*, pp.403–412, New York, NY, USA, ACM.

Hang, C., Wang, Y. and Singh, M. (2008) 'An adaptive probabilistic trust model and its evaluation', in *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multi Agent Systems*, International Foundation for Autonomous Agents and Multi agent Systems, Vol. 3, pp.1485–1488.

Hong, D. and Shen, V.Y. (2008) 'Setting access permission through transitive relationship in web based social networks', in *SWKM*.

Huang, B., Kimnig, A., Getoor, L. and Golback, L. (2013) 'A flexible framework for probabilistic models of social trust', In A.M. Greenberg, W.G. Kennedy and N.D. Bos (Eds.): *SBP 2013, LNCS*, Vol. 7812, pp.265–273.

Hussain, F.K. and Chang, E. (2007) 'An overview of the interpretations of trust and reputation', *The Third Advanced International Conference on Telecommunications*.

Huynh, T., Jennings, N. and Shadbolt, N. (2006) 'An integrated trust and reputation model for open multi-agent systems', *Journal of Autonomous Agents and Multi-Agent Systems*, Vol. 13, No. 2, pp.119–154.

Jamali, M. and Ester, M. (2011) 'Mining social network for recommendations, tutorial at ICDM', *International Conference on Data Mining*, December 11–14, Canada.

Jarecki, J., Meder, B. and Nelson, J. D. (2013) 'The assumption of class-conditional independence in category learning', *CogSci 2013 Proceedings*, pp.2650–2655, ISBN 978-0-9768318-9-1.

Jensen, F.V. (2001) *Bayesian Networks and Decision Graphs*, Springer Verlag, New York.

Jøsang A. (2001) 'A logic for uncertain probabilities', *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, Vol. 9, No. 3, pp.279–212.

Jøsang, A. (2007) 'Probabilistic logic under uncertainty', *Proceedings of Computing: The Australian Theory Symposium*, Vol. 65, Ballarat, Australia.

Josang, A. and Pope, S. (2005) 'Semantic constraints for trust transitivity', in *Proc. of Asia-pacific Conference on Conceptual Modelling*, Vol. 43, Australia.

Jøsang, A., Marsh, S. and Pope, S. (2006) 'Exploring different types of trust propagation', in Stolen, K. et al. (Eds.): *Lecture Notes in Computer Science*, Vol. 3986, pp.179–192, Springer Berlin Heidelberg.

Kazai, G. and Milic-Frayling, N. (2008) 'Trust, authority and popularity in social information retrieval', in *CIKM*, pp.1503–1504.

Keller, J.M., Gray, M.R. and Givnens Jr., J. A. (1985) 'A fuzzy K-nearest neighbour algorithm', *IEEE Transactions on Systems, Man, and Cybernetics*, July/August, Vol. SMC-15, No. 4, pp.580–585.

Keser, C. (2003) 'Experimental games for the design of reputation management systems', *IBM Systems Journal*, Vol. 42, No. 3, pp.498–506.

Khosravifar, B., Bentahar, J., Gomrokchi, M. and Alam, R. (2012) 'CRM: an efficient trust and reputation model for agent computing', *Journal of Knowledge-based Systems*, Vol. 30, pp.1–16, DOI: 10.1016/j.knosys.2011.01.004.

Kunegis, J., Lommatzsch, A. and Bauckhage, C. (2009) 'The Slashdot Zoo: mining a social network with negative edges', in *Proc. Int. World Wide Web Conf.*, pp.741–750.

Kuter, U. and Goldbeck, J. (2010) 'Using probabilistic confidence models for trust inference in web-based social networks', *ACM Trans. Internet Technol*., Vol. 10, No. 2, pp.1–23.

Laia, I.K., Tongb, V.W. and Laic, D.C. (2011) 'Trust factors influencing the adoption of internet-based inter organizational systems', *Journal of Electronic Commerce Research and Applications*, Vol. 10, No. 1, pp.85–93.

Leister, W. and Schulz, T. (2012) 'Ideas for a trust indicator in the IoT', *SMART 2012: The 1st International Conference on Smart System, Devices and Technologies*, *IARIA 2012*, pp.31–34.

Leskovec, J., Huttenloche, D. and Kleinberg, J. (2010) 'Predicting positive and negative links in online social networks', the *International World Wide Web Conference Committee (IW3C2) WWW 2010*, April 26–30, Raleigh, North Carolina, USA.

Levien, R. and Aiken, A. (1998) 'Attack-resistant trust metrics for public key certification', in *7th USENIX Security Symposium*, pp.229–242.

Liao, C., Liu, C. and Chen, K. (2011) 'Examining the impact of privacy, trust and risk perceptions beyond monetary transactions: an integrated model', *Journal of Electronic Commerce Research and Applications*, Vol. 10, No. 6, pp.702–716.

Massa, P. and Avesani, P. (2004) 'Trust-aware collaborative filtering for recommender systems', *Lecture Notes in Computer Science*, pp.492–508.

Massa, P. and Avesani, P. (2006) 'Trust-aware bootstrapping of recommender systems', *Proceedings of ECAI 2006 Workshop on Recommender Systems*, pp.29–33.

Massa, P. and Avesani, P. (2006) *Trust-aware Bootstrapping of Recommender Systems*, ECAI, Riva del Garda, Italy [online] http://www.gnuband.org/papers/trustaware_bootstrapping_of_recommender_ systems.

Massa, P. and Bhattacharjee, B. (2004) 'Using trust in recommender systems: an experimental analysis', *Trust Management: Second International Conference, ITrust 2004*, March 29–April 1, Springer, Oxford, UK.

McKnight, D.H. and Chervany, N.L. (2001) 'Trust and distrust definitions: one bite at a time', in Falcone, R., Singh, M. and Tan, Y.H. (Eds.): *Trust in Cyber-Societies: Integrating the Human and Artificial Perspectives*, pp.27–54, Springer, Berlin.

Mohanty, R., Ravi, V. and Patra, M.R. (2010) 'Web-services classification using intelligent techniques', *Expert Syst. Appl*., Vol. 37, No. 7, pp.5484–5490.

Mui, L., Mohtashemi, M. and Halberstadt, A. (2002) 'A computational model of trust and reputation', *Proceedings of the 35th Hawaii International Conference on System Science*.

Nefti, S., Meziane, F. and Kasiran, M.K. (2005) 'A fuzzy trust model for e-commerce', in *CEC*, pp.401–404.

Nitti, M., Girau, R. and Atzori, L. (2013) *Trustworthiness Management in Social IoT: First Theoretical Analysis*, PhD thesis, Veprints.umica.it/925/1/PhD_thesis_Nitti.pdf.

Patel, J., Luke Teacy, W.T., Jennings, N.R. and Luck, M. (2005) 'A probabilistic trust model for handling inaccurate reputation sources', in *iTrust*, pp.193–209.

Patel, J., Teacy, W., Jennings, N. and Luck, M. (2005) 'A probabilistic trust model for handling inaccurate reputation sources', *Third International Conference on Trust Management*, 23–26 May, Rocquencourt, France, pp.193–209.

Pearl, J. (1988) *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Ch. 5, pp.153–190, Morgan Kaufmann, California.

Qiu, X., Zhang, L., Wang, S. and Qian, G. (2010) 'A trust transitivity model based on Dempster-Shafer theory', *Journal of Networks*, Vol. 5, No. 2, pp.1025–1032.

Quercia, D., Hailes, S. and Capra, L. (2007) 'Lightweight distributed trust propagation', *Data Mining, IEEE International Conference on*, pp.282–291.

Quercia, D., Haites, S. and Capra, L. (2006) 'B-Trust: Bayesian trust framework for pervasive computing', in *Proc. of iTrust '06 International conference on Trust Management*, pp.298–312, Springer.

Rathore, A.S., Mutalikdesai, M.R. and Patil, S. (2013) 'Analysing trust based mixing patterns in signed networks', *15th International Conference on Asia-Pacific Digital Libraries*, *ICADL 2013*, Bangalore, India, December 9–11, *Proceedings, Lecture Notes in Computer Science*, Vol. 8279, pp.63–72.

Robinson, J., Wakeman, I., Chalmers, D. and Horsfall, B. (2010) 'Trust and the internet of things', in *TruLOCO: The Joint International Workshop on Trust Location and Communication Decentralized Computing*, Japan.

Santos-Neto, E., Ripeanu, M. and Iamnitchi, A. (2007) 'Tracking user attention in collaborative tagging communities', in *CAMA*, pp.11–18.

Shakeri, H. and Bafghi, A.G. (2014) 'A layer model of a confidence-aware trust management system', *International Journal of Information Science and Intelligent System*, Vol. 3, No. 1, pp.73–90.

Sillence, E., Briggs, P., Fishwick, L. and Harris, P.R. (2006) 'A framework for understanding trust factors in web-based health advice', *Journal of Human-Computer Studies*, Vol. 64, No. 8, pp.697–713.

Silver, M., Sakara, T., Su, H.C., Herman, C., Dolins, S.B. and O'shea, M.J. (2001) 'Case study: how to apply data mining techniques in a healthcare data warehouse', *Healthc. Inf. Manage*, Vol. 15, No. 2, pp.155–164.

Studeny, M. and Vomlel, J. (2011) 'On open questions in the geometric approach to structural learning Bayesian nets', Accepted in *International Journal of Approximate Reasoning*.

Toumi, A., Khenchaf, A. and Hoeltzener, B (2012) 'A retrieval system from inverse synthetic aperture radar images: application to radar target recognition', *Information Science*, Vol. 196, No. 1, pp.73–96.

Wang, Y., Lin, K-J., Wong, D.S. and Varadharajan, V. (2009) 'Trust management towards service-oriented applications', *Service Oriented Computing and Applications*, Vol. 3, No. 2, pp.129–146.

Witten, I.H. and Frank E. (2005) *Data Mining-Practical Machine Learning Tools and Techniques*, 2nd ed., Elsevier, UK.

Wua, J.J., Chenb, Y.H. and Chung, Y.S. (2010) 'Trust factors influencing virtual community members: a study of transaction communities', *Journal of Business Research*, Vol. 63, No. 9, pp.1025–1032.

Zadeh, L.A. (2002) 'From computing with numbers to computing with words from manipulation of measurements to manipulation of perceptions', *International Journal of Applied Math and Computer Science*, Vol. 12, No. 3, pp.307–324.

Zhang, X. and Zhang, Q. (2005) 'Online trust forming mechanism: approaches and an integrated model', *7th International Conference on Electronic Commerce*, pp.201–209.

Ziegler, C-N. and Lausen, G. (2004) 'Spreading activation models for trust propagation', in *Proceedings of the IEEE International Conference on e-Technology, e-Commerce, and e-Service*, IEEE Computer Society Press, Taipei, Taiwan, March [online] http://www.citeseer.ist.psu.edu/ziegler04spreading.html.