Can reputation migrate? On the propagation of reputation in multi-context communities

Bagheri, Ebrahim; Zafarani, Reza; Barouni-Ebrahimi, M.

Publisher's version / Version de l'éditeur:
https://doi.org/10.1016/j.knosys.2009.05.007
Knowledge-Based Systems Journal, 22, 6, pp. 410-420, 2009-06-02

Can Reputation Migrate? On the Propagation of Reputation in Multi-Context Communities

Ebrahim Bagheri, Reza Zafarani, and M. Barouni-Ebrahimi

Faculty of Computer Science,
University of New Brunswick, Fredericton, Canada
{e.bagheri,r.zafarani,m.barouni}@unb.ca

Abstract. As e-communities grow in both quality and quantity, their online users require more appropriate tools to suite their needs in such environments. Many such tools are not explicitly needed in real-world communities where humans directly interact with each other. Trust making and reputation ascription are among the most important examples of such tools. Humans often build trust relationships through interaction or recommendation, and are therefore able to ascribe relevant reputation to those they interact with. However, in online communities the process of trust making and reputation ascription is more complicated. In this paper, we address a special case of the trust making process where community users need to create bonds with those they have not encountered before. This is a common situation in websites such as amazon.com, ebay.com, epionions.com and many others. The model we propose is able to estimate the possible reputation of a given identity in a any new context by observing his/her behavior in other communities. Our proposed model employs Dempster-Shafer based valuation networks to develop a global reputation structure and performs a belief propagation technique to infer contextual reputation values. The preliminary evaluation of the proposed model on a dataset collected from epinions.com shows promising results.

1 Introduction

Reputation is a distributed, socially ascribed, and collective belief of a society towards the stand-point of a single person, group, role or even a non-human entity within the context of that given society [20]. Reputation is developed based on the general belief of society actors whether or not a given identity has fully satisfied the expectations of its roles. If an actor fails to comply with social standards and show acceptable performance, it will develop a negative reputation amongst the rest of the community members. Similarly, if the society develops a positive perception of that identity's effectiveness, it will be rewarded with the attribution of a positive reputation [4]. A high degree of reputation directly contributes to the development of stronger social status and influence; however, a weak or even negative reputation discourages society members to develop the required trust for embarking on any social interaction with the reputation holder

[13]. It is hence to the participators high interest to optimize their performance to construct a positive social face.

The notion of reputation has been employed in various application areas, such as electronic market places [19], peer to peer systems [18], information sharing communities [4], and recommender systems [3], to name a few. It is believed that the formalization of reputation is a context-dependent process. This means that the formal definition of reputation structure is reliant on contextual features, societal values, and environmental goals of the target domain where reputation is being defined and deployed. As an example, Carter et al [4] have employed five properties of an information sharing community, namely social information provisioning, degree of interaction, effort for content preparation, role in environment administration, and longevity to formalize the concept of reputation in a knowledge-exchange society. In such an environment, the involvement of a community member in social information provisioning is an important factor that has a synergistic effect on positive reputation. In contrast, in an environment where critical information is being secured and disclosure is undesirable, sharing of information can negatively affect one's reputation. This situation clearly shows that the constituting elements of reputation are either different or their behavior is dissimilar under unrelated contexts. Such cases necessitate the development of a structured framework for handling reputation in environments where multiple contexts coexist.

Even though the assessment and ascription of reputation is fundamentally a process that relies on specific contextual values and norms, it is common practice within human societies, to consider one's reputation in a certain context to infer his/her reputation in other contexts. For instance, people with high reputation values in a particular context will be looked upon as successful and highly reputed members of the society even from the perspective of other contexts. The opposite situation may also be true where people with low reputation values in a certain context will not be highly appreciated due to their low reputation in other contexts. This shows that while the formalization of reputation is to a great extent context-dependent, still the mutual effects of different contexts on each other cannot be overlooked.

As mentioned in both Kinateder et al. [8] and Seigneur et al [15], there are some intricate interdependencies between personal trust values of different contexts. This implicitly states that reputation values in various contexts can have indirect influence on each others' formation. A clear example given by Seigneur et al. [15] argues that for instance, a skilled cook that has previously committed homicide may not be considered as a trustworthy (reputable) cook. His/her low social status and lack of trustworthiness developed as a result of the homicide can result in a low reputation for him/her in other contexts such as his/her job as a chef. It is however important to note here that one's reputation in a given context does not always enjoy a direct relation with reputation in other contexts and may have an inverse effect. For instance, generally prominent students in mathematics are not good sportsmen and vice versa!

The independence of reputation formalization in different contexts on the one hand, and the obscure implicit relationship and effect of reputation in conceptually related contexts on the other hand, complicates the development of a global reputation value for each entity in multi-context environments. For example, evaluating the position of a person in a certain context where he/she has not had any previous self presentation becomes a challenging task. In real-world situations, people carry some extent of their reputation with them into new contexts. Although the initial reputation that they carry with them is not an exact determination of the ultimate reputation that will be ascribed in the new context (since reputation is dynamic and changes based on the person's behavior in social processes), it can be regarded as a good estimation of the person's initial standing point. The person can further improve or damage his initial reputation through social interactions within the framework of the new context.

Previously, we have proposed a framework for dynamically updating and inferring the unobserved reputation of environment participants in different contexts [2] . This framework proposes the employment of a reputation structure tree to represent the relationship between the contexts of the environment. Reputation of a given identity in one context can be propagated to other contexts through two mechanisms, namely: *forward update* and *backward adjustment*. The advantage of this model is that the propagation of reputation has been shown to reach equilibrium after regular changes in contextual reputations. The shortcoming of this work is that it does not provide suitable means for developing the reputation structure tree, and its performance has not been evaluated on real-world data.

Aberer and his colleagues state that the current state of the art in reputation management systems can be categorized in two main classes: 1) those which employ social networking features by accumulating all of the available feedback in the community in order to develop a robustness reputation estimation mechanism; and 2) probabilistic methods that rely on probabilistic estimation techniques on a limited fraction of the available information in the community [6, 17, 5]. The model that we propose in this paper relies on both approaches, by employing the concept of *valuation networks* [16]. In the proposed model, global reputation is modeled as Dempster-Shafer belief functions on a *Markov tree* through which the relationship between various contexts of a unique environment are modeled through hyper-vertices of the Markov tree. Reputation of each identity in a given context is represented using a *belief mass assignment* function. The estimation of reputation in various contexts of the environment is performed by the employment of the message passing-based belief propagation model of the Shenoy-Shafer architecture [14].

It should be noted that the model proposed in this paper is quite unique and different from the current reputation management systems in that the proposed model does not engage in the process of deriving, calculating, updating or storing reputation values. The basic assumption of our system is that an established reputation inference system is already used for calculating contextual reputation, and therefore, it only focuses on the propagation of already observed contextual reputation to unobserved contexts [1]. Several available reputation calculation systems already exist in the related literature that can be used to derive contextual reputation [11, 4, 13, 18].

The remainder of this paper is structured as follows: in the next section, some preliminaries are introduced. A structured problem definition and the organization of the proposed model is given in Section 3. The application of the proposed model to the epinions.com website is studied in Section 4. The paper is then concluded in Section 5.

2 Preliminaries

We define Θ_x as the state space of the variable x. Each variable x should have a finite state space that is all the possible values for x are known. Given a set of variables denoted D, we let Θ_D represent the Cartesian product $\Theta_D = \times\{\Theta_x : x \in D\}$. Θ_D is called the state space of D; therefore, the members of Θ_D are configurations of D.

2.1 Basics of Dempster-Shafer Theory of Evidence

Dempster-Shafer (DS) theory of evidence is one of the most widely used models that provides means for approximate and collective reasoning under uncertainty [7]. It is basically an extension to probability theory where probabilities are assigned to sets as opposed to singleton elements. The employment of the DS theory requires the definition of the set of all possible states in a given setting, referred to as the frame of discernment represented by θ_D. The powerset of θ_D, denoted 2^{θ_D}, incorporates all possible unions of the sets in θ_D that can receive belief mass.

The truthful subsets of the powerset can receive a degree of belief mass; therefore, the belief mass assigned to an atomic set such as $\psi \in 2^{\theta_D}$ is taken as the belief that the given set is true. Moreover, the belief mass ascribed to a non-atomic set such as $\psi \in 2^{\theta_D}$ is interpreted as the belief that one of the atomic sets in ψ is true, but uncertainty rules out the possibility of pinpointing the exact atomic set.

Definition 1 *A belief mass assignment function is a mapping* $[\phi]_m : 2^{\theta_D} \rightarrow [0, 1]$ *that assigns* $[\phi(A)]_m$ *to each subset* $A \in 2^{\theta_D}$ *such that:*

$$[\phi(A)]_m \geq 0, \tag{1}$$

$$[\phi(\varnothing)]_m = 0, \tag{2}$$

$$\sum_{A \in 2^{\theta}} [\phi(A)]_m = 1. \tag{3}$$

The belief in A is interpreted as the absolute faith in the truthfulness of A, which not only relies on the belief mass assigned to A but also to belief masses assigned to subsets of A.

Definition 2 *A belief function corresponding with $[\phi(B)]_m$, a belief mass assignment on θ_D, is a function $[\phi]_b : 2^{\theta_D} \rightarrow [0,1]$ defined as:*

$$[\phi(B)]_b = \sum_{A \subseteq B} [\phi(A)]_m, \quad A, B \in 2^{\theta_D}. \tag{4}$$

For any belief functions $[\phi]_b$ defined over D, D is called the domain of $[\phi]_b$. All subsets of Θ_D ($A \subseteq \Theta_D$) that satisfy $[\phi(A)]_m \neq 0$ are known as focal sets of Θ_D.

The fundamental operations of the Dempster-Shafer theory are the *combination* and the *marginalization* functions. These operations allow aggregation and coarsening, respectively. Marginalization takes a mass assignment function $[\phi(A)]_m$ on domain D and produces a new mass assignment function $[\phi'(A)]_m$ on domain $C \subseteq D$.

Definition 3 *Let $[\phi(A)]_m$ be a mass assignment function on domain D and let $C \subseteq D$. The marginalization of $[\phi(A)]_b$ to C produces a belief function over C:*

$$[\phi^{\downarrow C}(B)]_m = \sum_{A: A^{\downarrow C} = B} [\phi(A)]_m. \tag{5}$$

The base combination rule for multiple mass assignment functions within the framework of Dempster-Shafer theory of evidence is Dempster's rule of combination, which is a generalization of Bayes' rule [12]. This combination operator ignores the conflicts between the functions and emphasizes on their agreements.

Definition 4 *Let $[\phi_1(A)]_m$ and $[\phi_2(A)]_m$ be two mass assignment functions on domain D_1 and D_2, respectively. The non-normalized combination of these two functions produces a mass assignment function over $D = D_1 \cup D_2$:*

$$[\phi_1(A)]_m \oplus [\phi_2(A)]_m = \sum_{B_1, B2} \{[\phi_1(B_1)]_m \cdot [\phi_2(B_2)]_m \quad : \quad B_1^{\uparrow D} \cap B_2^{\uparrow D} = A\}, \tag{6}$$

where $B^{\uparrow D}$ represents B extended to domain D. The normalized form of the combination operator is defined as:

$$[\phi_1(A)]_m \oplus [\phi_2(A)]_m = \frac{\sum_{B_1, B2} \{[\phi_1(B_1)]_m \cdot [\phi_2(B_2)]_m : B_1^{\uparrow D} \cap B_2^{\uparrow D} = A\}}{1 - \sum_{B_1, B2} \{[\phi_1(B_1)]_m \cdot [\phi_2(B_2)]_m : B_1^{\uparrow D} \cap B_2^{\uparrow D} = \varnothing\}}. \tag{7}$$

2.2 Belief Propagation

The corresponding domains of a set of mass assignment functions $[\phi_1]_m, [\phi_2]_m, ..., [\phi_n]_m$ form a hypergraph. From the hypergraph, a covering hypertree can be developed that can be employed to construct a Markov tree.

Definition 5 *Let $(\mathcal{V}, \mathcal{E})$ be a tree where $\mathcal{V}$ is the set of vertices, and $\mathcal{E}$ is the set of its edges. Each $v \in \mathcal{V}$ is itself a non-empty set; therefore, $\forall v \in \mathcal{V}$, v is a hypergraph. We call $(\mathcal{V}, \mathcal{E})$ a hypertree if the following conditions are satisfied:*

(1) $\mathcal{V}$ *is a hypertree,*
(2) if $(\{v, v'\} \in \mathcal{E})$ *then* $v \cap v' \neq \varnothing$*,*
(3) if v *and* v' *are distinct vertices, and* X *is in both* v *and* v'*, then* X *is in every vertex on the path from* v *to* v'*.*

Each node of the Markov tree consists of a belief mass assignment function $[\phi_i]_m$ that operates over Domain D_i. Belief propagation in the Markov tree in the Shenoy-Shafer architecture is performed through a message passing scheme [14]; hence, each node in the Markov tree will possess the global mass assignment function $[\phi]_m = ([\phi_1]_m \oplus \cdots \oplus [\phi_n]_m)$ marginalized to its own domain D_i. In order to perform message passing, and enable local computation, two base operations: *marginalization* and *combination*, and three simple axioms, namely: *transitivity of marginalization*, *commutativity and associativity of combination*, and *distributivity of marginalization over combination* are required [9].

Belief propagation over a Markov tree structure (valuation network) can be performed by passing messages between tree nodes using the following two rules:

1) Each node sends a message to its neighbor. Let $\mu^{A \to B}$ be a message from A to B, and $\Phi(A)$ be the set of neighbors of A and the belief mass assignment of A be $[\phi_i(A)]_m$, then the passed message from A to B is defined as a combination of messages from all neighbors of A except B and the belief mass of A:
$$\mu^{A \to B} = \left(\oplus\{\mu^{X \to A} | X \in (\Phi(A) - \{B\}) \oplus [\phi_i(A)]_m\}\right)^{\downarrow A \cap B} \qquad (8)$$
2) When node B receives messages from its neighbors, it combines all the received messages with its own belief mass and employs the result as its own marginal.

3 Model Overview

The model we propose in this paper, will formally address the following two issues:

1) The definition of a structure for permitting global reputation management through local-contextual reputation computations,
2) The development of a theoretical framework for the propagation of local-contextual reputation values as belief masses between different contexts.

In Section 1, we informally mentioned that there are two main characteristics for multi-context reputation formalization that complicates the definition of a straightforward reputation estimation model, i.e. independence of reputation formalization in different contexts, and implicit relationship between reputation in conceptually related contexts. Here, we show how the reputation of various contexts of an environment can have implicit relationship and cross-context impact, while preserving their independence by definition. We first define the interpretation of contextual reputation.

Definition 6 *The interpretation of a contextual reputation $\mathcal{CR}_i$ for context i represented through a set of social (contextual) norms, denoted $\mathcal{N}$, is a pair (U, I) where U is the domain of interpretation of all contextual reputations $\mathcal{CR}$, and I is a total morphism that maps $\forall \mathcal{N}_i \in \mathcal{N}$ onto a relation R.*

An interpretation of a contextual reputation relates each social (contextual) norm onto its corresponding underlying concept in the contextual domain of discourse; therefore, an interpretation conveys how the contextual reputation is understood with regards to the given domain of discourse based on context foundations.

Definition 7 *Let $\mathcal{N}_{i,k}$ and $\mathcal{N}_{j,l}$ be only two social norms of two different contextual reputation formalizations $\mathcal{CR}_i$, and $\mathcal{CR}_j$, and let $T_i = (U_i, I_i)$, and $T_j = (U_j, I_j)$ be interpretations of $\mathcal{CR}_i$ and $\mathcal{CR}_j$, respectively. The implicit impact relationships between $\mathcal{CR}_i$ and $\mathcal{CR}_j$ can be defined as follows: With $I_i(\mathcal{N}_{i,k}) \neq \varnothing,\ I_j(\mathcal{N}_{j,l}) \neq \varnothing$:*

- *(no impact)* $I_i(\mathcal{N}_{i,k}) \cap I_j(\mathcal{N}_{j,l}) = \varnothing$,
- *(total impact)* $I_i(\mathcal{N}_{i,k}) = I_j(\mathcal{N}_{j,l})$,
- *(inclusive impact)* $I_i(\mathcal{N}_{i,k}) \subseteq I_j(\mathcal{N}_{j,l})$,
- *(partial impact)* $I_i(\mathcal{N}_{i,k}) \cap I_j(\mathcal{N}_{j,l}) \neq \varnothing,\ I_i(\mathcal{N}_{i,k}) - I_j(\mathcal{N}_{j,l}) \neq \varnothing, I_j(\mathcal{N}_{j,l}) - I_i(\mathcal{N}_{i,k}) \neq \varnothing$.

Definition 7 shows that two contextual reputations can only have implicit impact on each others behavior if the interpretation of their social norms have some degree of overlap; therefore, contexts with no overlapping social norms have no degree of effect on each other, while contexts with total overlapping social norm interpretations have total impact on each other. The degree of impact, denoted $\mathcal{DI}$, can be easily defined using the interpretation of contextual norms as:

$$\mathcal{DI}(I_i, I_j) = \frac{I_i(\mathcal{N}_i) \cap I_j(\mathcal{N}_j)}{I_i(\mathcal{N}_i) \cup I_j(\mathcal{N}_j)}. \tag{9}$$

Based on Definitions 6 and 7 global reputation for a specific environment comprising multiple contexts can be defined as a hypergraph.

Definition 8 *Let $\mathcal{EN}$ be an environment, and $\mathcal{C}_{\mathcal{EN}}$ be the set of all its contexts. Global reputation $\mathcal{GR}$ for $\mathcal{EN}$ is defined as a hypergraph $(\mathcal{V}, \mathcal{E})$ where $\mathcal{V}$ is the set of all its vertices, and $\mathcal{E}$ all of its edges, such that $\mathcal{V} \equiv \mathcal{C}_{\mathcal{EN}}$, and $\forall \mathcal{C}_i, \mathcal{C}_j$ that $\mathcal{DI}(I_i, I_j) > 0$ then $\{\mathcal{C}_i, \mathcal{C}_j\} \in \mathcal{E}$.*

The global reputation hypergraph $\mathcal{GR}$ can then be reduced to a Markov tree to avoid loops while reputation is propagated between its contexts. The reduced hypergraph converted into a Markov tree is the final structure for global reputation representation denoted $\mathcal{GR}^m$.

Each hyperedge of the $\mathcal{GR}^m$ Markov tree contains the global reputation value marginalized to its constituting vertices (contexts), i.e. the reputation of a given entity can be calculated for the contexts in that hyperedge using the belief mass assignments of this hyperedge through marginalization. Suppose that $\mathcal{GR}^m$ only contains two contexts (vertices): $\mathcal{C}^1_{\mathcal{EN}}$, and $\mathcal{C}^2_{\mathcal{EN}}$, and only one hyperedge $\mathcal{E}_1 = \{\mathcal{C}^1_{\mathcal{EN}}, \mathcal{C}^2_{\mathcal{EN}}\}$. The reputation of a given entity can be calculated for any of the two contexts by marginalizing the belief mass assignments of $\mathcal{E}_1$ onto that specific context. For instance, a reputation in $\mathcal{C}^1_{\mathcal{EN}}$ can be calculated through $[\phi_{\mathcal{E}_1}^{\downarrow \mathcal{C}^1_{\mathcal{EN}}}]_m$, where $[\phi_{\mathcal{E}_1}]_m$ represents the belief mass assignment of $\mathcal{E}_1$.

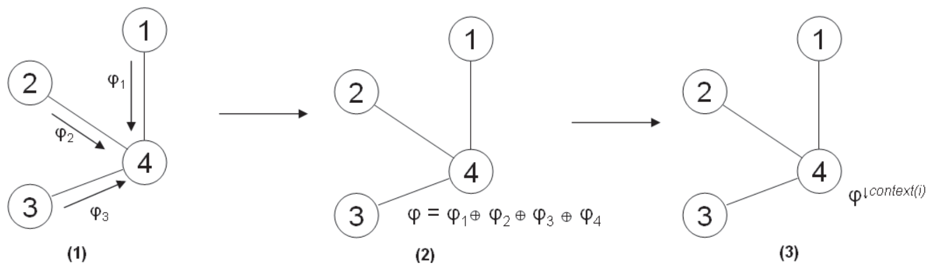

Fig. 1. Reputation Calculation for *Context* i in a Global Reputation Markov Tree ($\mathcal{GR}^m$)

In cases where $\mathcal{EN}$ consists of more than two contexts, and therefore, more than one hyperedge may exist, the process of local reputation computation over $\mathcal{GR}^m$ is more complex. In such a situation, the local belief mass assignments in each hyperedge of $\mathcal{GR}^m$ needs to be propagated to the target hyperedge (the hyperedge that requires the local-contextual reputation calculation) using the belief propagation scheme introduced in Section 2.2. Having received all the messages, the destination hyperedge that needs the computation of the local reputation value should compile all the messages into one belief mass assignment using Dempster's rule of combination. The resulting belief mass assignment can then be marginalized to the context of interest in order to calculate the local-contextual reputation. Figure 1 depicts the steps of this process.

In the following section, our experience in applying the proposed global reputation markov tree structure and the reputation propagation theme to compute and propagate reputation in the data collected from the epinions.com dataset is elaborated.

4 Experience with Epinions.com

For experimental purposes, we collected data from a popular online community: epinions.com. Epinions.com is a website that collects consumer experience reports and beliefs about various products. The range of products in this website covers a vast variety of different subjects conceptually grouped in hierarchical categories. The users are able to participate in epinions.com by writing reviews

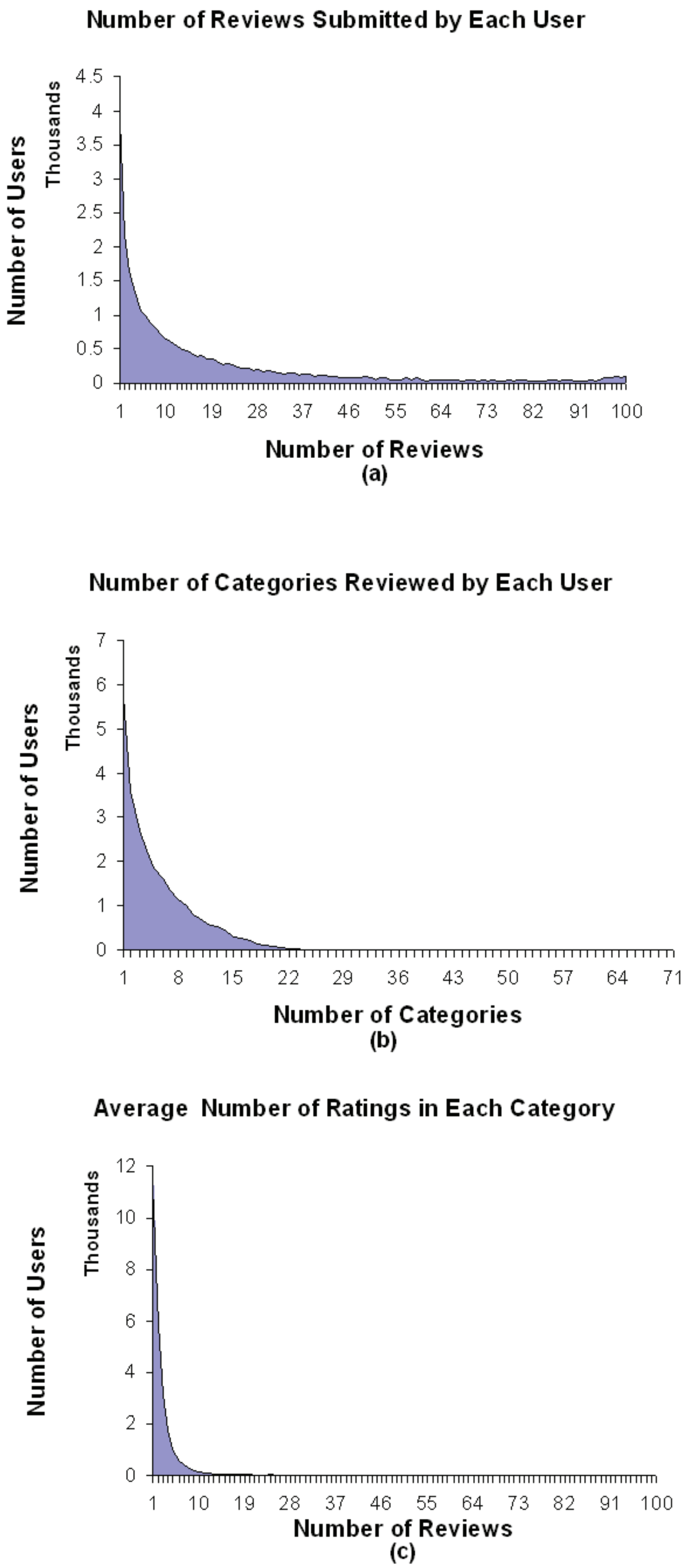

Fig. 2. The Behavior of the epinions.com Dataset.

about the products of different categories, rating the items, and also rating the previously written reviews (ratings are numeric values ranging from the minimum of 1 to a maximum of 5 - Table 1 shows a compilation of the extracted data). We collected information for $33,876$ distinct users from which only $25,541$ users had at least provided one review in the website. For each of the users up to 100 of their submitted reviews were also gathered. These reviews were from 71 different categories. Overall, $463,398$ reviews were accumulated. The dataset was structured as $\langle \mathcal{U}, \mathcal{C}, \mathcal{R} \rangle$, where $\mathcal{U}$, $\mathcal{C}$, and $\mathcal{R}$ represent the username, the category where the review was submitted, and the overall rating that the given review had received from the community, respectively.

The analysis of the collected dataset is complex, since most of the users in the epionions.com dataset suffer from the *cold start problem*, which is the submission of a very low number of reviews [10]. Figure 2 shows that the majority of the users have submitted less than 10 reviews, and that their participation is limited on average to less than 8 categories from the set of 71 possible categories. Furthermore, it can be seen that the average number of reviews written by each user for a given category is less than 7. Comparing these statistics, we can trivially infer that since an average user typically submits around 10 reviews in epinions.com and 7 reviews for a category where he/she frequently participates, that the users of epinions.com are mostly concentrated on one specific category of their interest. For instance, although a user specializing in Music submits several ratings to categories such as Movie, Travel, Sport, etc., his/her main contribution is towards the Music category. This issue makes the detection of conceptually related categories from the participation behavior of the users difficult.

In the following, assuming that each category of epinions.com is a context within the the online community environment, we will explain how the *degree of impact*, $\mathcal{DI}$, for any two categories is calculated. The formalization of $\mathcal{DI}$ would provide the basis for crafting the global reputation markov tree structure for the categories.

4.1 Formalizing the Degree of Impact ($\mathcal{DI}$)

Let us informally assume that two categories of an online community have some shared underlying principles which make them conceptually equally appealing for the community if there exist a common group of like-minded users in both of those categories. Based on this assumption, suppose there is a group of users $\mathcal{U}_c = \{u_1, ..., u_m\}$ that have reviewed items in category c, and each user has received $\mathcal{R}_{u,c} = \{r^u_{1,c}, ..., r^u_{n,c}\}$ ratings for his/her reviews in category c, while each user has submitted $\mathcal{N}_{u,c}$ reviews for that specific category. For any given two categories, we would believe them to be conceptually close, if the behavior of their shared users are similar, i.e. the number of ratings that each user submits in these two categories are alike ($\mathcal{N}_{u,c} \sim \mathcal{N}_{u,c'}$) and also the average rating received by that specific user is also similar in both categories ($\bar{r}_{i,c} \sim \bar{r}_{i,c'}$). Simply stated, we are looking for categories that have users with similar participation rate and review quality.

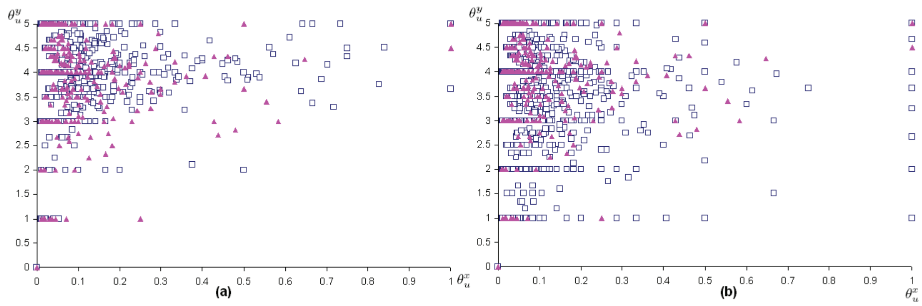

Fig. 3. The Non-normalized Overlapping Distribution of Data Points for Different Categories: (Movie) and (Home and Garden) in (a), and (Electronics) and (Computer Hardware) in (b).

For each category $\mathcal{C}_i$, and the set of users participating in $\mathcal{C}_i$, denoted $\mathcal{U}_{c_i}$, we can form a set of data points $\Theta_{\mathcal{C},i} = \{\theta_{\mathcal{C}_i,1}, ..., \theta_{\mathcal{C}_i,n}\}$, where $\theta_{\mathcal{C}_i,u}$ for a given user $u \in \mathcal{U}_{c_i}$ is defined as follows:

$$\theta_{\mathcal{C}_i,u} = (\theta^x_{\mathcal{C}_i,u}, \theta^y_{\mathcal{C}_i,u}). \tag{10}$$

$$\theta^x_{\mathcal{C}_i,u} = \frac{\mathcal{N}_{u,\mathcal{C}_i}}{\sum_{c\in\mathcal{C}} \mathcal{N}_{u,c_i}}. \tag{11}$$

$$\theta^y_{c_i,u} = \frac{\sum_{j=1}^{\mathcal{N}_{u,c_i}} r^u_{j,c_i}}{\mathcal{N}_{u,c_i}}. \tag{12}$$

where $\theta^x_{\mathcal{C}_i,u}$ represents the fraction of effort that a given user u has spent in category $\mathcal{C}_i$, and $\theta^y_{\mathcal{C}_i,u}$ shows the average quality of the reviews written by user u in category $\mathcal{C}_i$ (ascribed by other community members in the form of ratings).

Figure 3 shows the overlapping distribution of the data points ($\theta_{\mathcal{C}_i,u}$) of categories: (Movie) and (Home and Garden) in (a), and (Electronics) and (Computer Hardware) in (b), respectively. As was observed in Figure 3, the overlap of two distributions for any two categories produces a similar non-separable overlap pattern, which is undesirable when high discriminative power is required to distinguish between high impact and low impact categories.

To discriminate between the closely related and poorly associated categories, we perform an iterative normalization process on the data points of each category before comparing them. In each iteration the data points of the previous iteration are normalized and $\tilde{\Theta}^{i+1}_{\mathcal{C}_j} = \{\frac{\tilde{\theta}^i_{\mathcal{C}_j,1} - \mu_{\tilde{\Theta}^i_{\mathcal{C}_j}}}{\sigma_{\tilde{\Theta}^i_{\mathcal{C}_j}}}, ..., \frac{\tilde{\theta}^i_{\mathcal{C}_j,n} - \mu_{\tilde{\Theta}^i_{\mathcal{C}_j}}}{\sigma_{\tilde{\Theta}^i_{\mathcal{C}_j}}}\}$ is developed. This process is continued until $|\mu_{\tilde{\Theta}^{i+1}_{\mathcal{C}_j}} - \mu_{\tilde{\Theta}^i_{\mathcal{C}_j}}| < \epsilon$ where $\mu_{\tilde{\Theta}^{i+1}_{\mathcal{C}_j}}$ and $\mu_{\tilde{\Theta}^i_{\mathcal{C}_j}}$ represent the average of values in $\tilde{\Theta}^i_{\mathcal{C}_j}$ and $\tilde{\Theta}^i_{\mathcal{C}_j}$, respectively. The result of this process can be seen in Figure 4 where in (a), the data points have been clearly separated (the majority of data points on the right are non-overlapping) while in (b) the

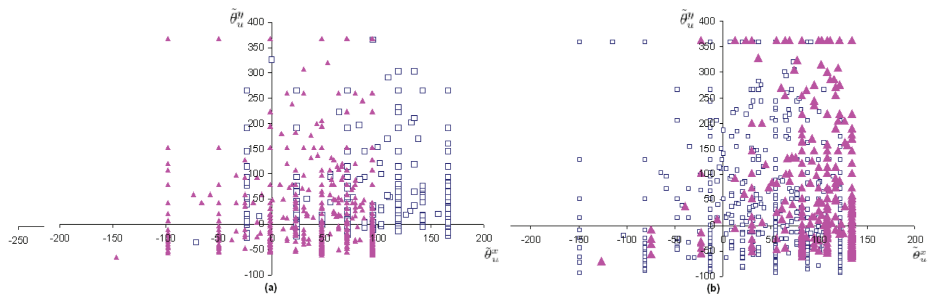

Fig. 4. The Normalized Overlapping Distribution of Data Points for Different Categories: (Movie) and (Home and Garden) in (a), and (Electronics) and (Computer Hardware) in (b).

data points are totally overlapping, whereas this discrimination could not have been observed in Figure 3. The degree of overlap on these data distributions can show the degree of relative closeness of any two categories; therefore, it can be inferred from Figure 4 that categories (Electronics) and (Computer Hardware) are more closely related as compared to (Movie) and (Home and Garden). Let us now define the *degree of impact* ($\mathcal{DI}$) as the Euclidean distance of the centers of the distributions of each two categories; therefore, let $\wp_i$ and $\wp_j$ be the centers of the data point distributions of categories $\mathcal{C}_i$, and $\mathcal{C}_j$, the degree of impact of these two categories on each other, denoted $\mathcal{DI}_{\mathcal{C}_i,\mathcal{C}_j}$, is defined as:

$$\mathcal{DI}_{\mathcal{C}_i,\mathcal{C}_j} = \left(|\wp_i^x - \wp_j^x|^2 + |\wp_i^y - \wp_j^y|^2\right)^{1/2}. \tag{13}$$

It is important to notice that most commonly used methods such as Latent Semantic Indexing (LSI) that employ Singular Value Decomposition (SVD) are not applicable to the epinions.com dataset or other similar datasets due to some of the specific characteristics of these datasets: First of all, the high number of users and the corresponding submitted ratings introduces a high computational complexity in the singular value decomposition process (which is of order $O(min(nm^2, mn^2))$ where m and n are the matrix dimensions). Secondly, due to the high sparsity of user-category rating matrix, the matrix is nearly full-rank and therefore the rank lowering process of SVD is more or less redundant. Therefore, such methods are unable to find the distance and correspondences of community contexts with each other.

4.2 Forming the Global Reputation Markov Tree Structure

The degree of impact of two categories can be interpreted as a distance measure between two categories. With this interpretation, a fully connected weighted graph represented by an $n \times n$ matrix, $G_{\mathcal{EN}}$, where n is the number of categories in $\mathcal{EN}$, can be created such that:

$$G_{\mathcal{EN}}(\mathcal{C}_i, \mathcal{C}_i) = \infty, \tag{14}$$

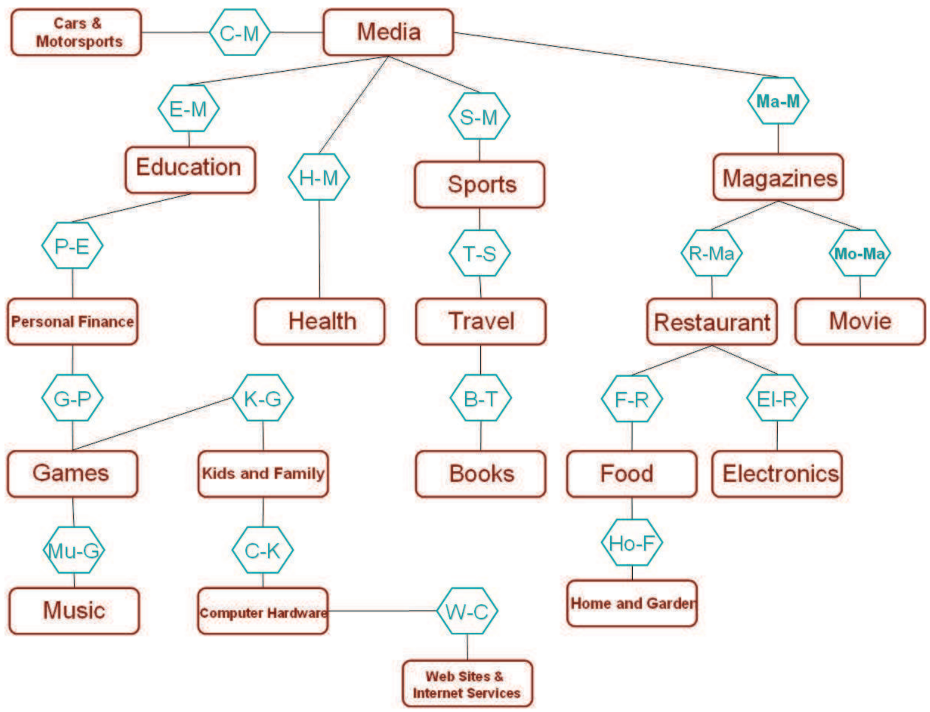

Fig. 5. The Global Reputation Markov Tree Structure

$$G_{\mathcal{EN}}(\mathcal{C}_i, \mathcal{C}_j) = \mathcal{DI}_{\mathcal{C}_i,\mathcal{C}_j}. \tag{15}$$

Kruskal's algorithm was further applied on the formed matrix $G_{\mathcal{EN}}$, to find the minimum spanning tree ($\tilde{G}_{\mathcal{EN}}$) for the corresponding connected weighted graph. The neighboring nodes of $\tilde{G}_{\mathcal{EN}}$ with equal distances were then cloned into a single node to form a hypertree required for the belief propagation scheme. The final result shown in Figure 5 illustrates the global reputation Markov tree structure developed by only considering the top 19 most active categories in the epinions.com dataset.

4.3 Allocating Initial Reputation Belief Masses

Each node of the global reputation Markov tree contains a joint belief distribution over the categories that form it. For instance, in Figure 5, the belief mass assignment in the node $C - M$ is distributed over the categories: (Media) and (Cars & Motorsports). In order to be able to propagate belief through the reputation tree, the belief mass assignment function of each node should be specified. The simplest way to do this is to initially assign the highest possible belief mass to the superset, and gradually, as new evidence about the performance of a given user in the environment is received, the belief mass assignments of that specific user is updated to reflect the information gained from his/her performance.

It is also possible to assign an initially similar belief mass assignment for all the users based on the information in the epinions.com dataset, and then update that belief mass assignments for each user separately. The initial belief mass assignment should assign belief masses based on the frequency of observation of the possible states related to the categories of each node. Let's suppose that reputation is described using High, Medium, and Low values. In our experiments, we suppose that the average rating that a user's reviews in a specific category receives represents the user's reputation in that category. We convert the five scale ratings $[1, 5]$ into High, Medium, and Low. Therefore as an example, the state space which can receive belief mass for the $C - M$ node would be $\Theta_{C-M} = \{(High, High), (High, Med), (High, Low), (Med, High), (Med, Med), (Med, Low), (Low, High), (Low, Med), (Low, Low)\}$. The belief mass assigned to a state such as (High,High) for a node with two categories: $\mathcal{C}_i$ and $\mathcal{C}_j$ is calculated as follows:

$$[\phi_{(\mathcal{C}_i,\mathcal{C}_j)}(High, High)]_m = \frac{\mathcal{N}(High, High)}{\sum_{\mathcal{X},\mathcal{Y}\in\mathcal{RP}} \mathcal{N}(\mathcal{X}, \mathcal{Y})} \tag{16}$$

where $\mathcal{RP} = \{High, Medium, Low\}$, and $\mathcal{N}(High, High)$ is the number of users that have a High reputation in $\mathcal{C}_i$ and also a High reputation in $\mathcal{C}_j$. It is clear that this initial belief mass assignment model respects the conditions of Definition 1. Now with this formula, the initial belief mass assignment of all nodes of the global reputation tree can be easily calculated.

5 Performance Evaluation

In this section, a comprehensive evaluation of our model is presented. The generalization capabilities of the model as well as three critical properties of it are thoroughly assessed. First, the generalization evaluation of the model in terms of its accuracy in predicting hidden reputations is presented. Secondly, the expected similarity and the average similarity values of the model are measured under different situations in which the user, domain, or both are unknown.

5.1 Reputation Estimation Analysis

In the first set of experiments, the collected epinions.com dataset was split into two testing and training datasets for evaluation purposes. The training dataset was employed to create the global reputation Markov tree structure and form the initial belief mass assignments of each node of the Markov tree. Based on the structure of the reputation tree we conducted an experiment similar to label prediction in machine learning datasets. In our experiments, for each user in the training dataset, we removed one of his/her known reputation values for a given category (one of the environments contexts), and tried to estimate the hidden reputation value based on the other available reputation values of the user in other categories.

The available reputation values for each user were handled as incoming evidences about the overall performance of the user in the other categories. These evidences were then propagated through the reputation structure towards the node containing the target category using the introduced belief propagation scheme, where the belief masses were combined and marginalized to that category. Finally, the state space (either High, Medium or Low) with the highest belief mass assigned to it was selected to serve as the representative of that users reputation in the given category. The proposed model's estimation capability showed to be able to correctly predict the missing reputation values in 76.923% of the test cases. In our opinion, this is a significant achievement, since a lucky guess by the other users of the target context to ascribe reputation and trust to a given user is only 33.3% successful. Our model increases the chance of correct reputation ascription in a given context where a user has not had any previous performance to an extent of 2.31 times.

5.2 Expected Reputation Similarity Analysis

To complement the results obtained in the previous set of experiments, an analysis of our technique is provided in the following three subsections. In these analyses, the overall reputation prediction accuracy, with respect to different users and categories is measured. In the beginning, a set of 30 random users were selected based on whose profile a sequence of experiments were conducted. First, a category was chosen from a user's profile to serve as his/her area of specialty. This category, which we call the user's *domain* hereafter, provides grounds to approximate the user's reputation in other categories. After the *domain* is determined, for all other categories, the user's reputation is estimated and compared to the real user's reputation in that category. Figure 6 depicts the similarity between the estimates and real reputation values, for a specific user and *domain*, over all categories.

To be more specific, the estimated reputation is represented using the following vector:

$$\hat{\mathcal{R}} :< \hat{P}(C_{i,h}|user, domain), \hat{P}(C_{i,m}|user, domain), \hat{P}(C_{i,l}|user, domain) > \tag{17}$$

where $\hat{P}(C_{i,h}|user, domain)$, $\hat{P}(C_{i,m}|user, domain)$, and $\hat{P}(C_{i,l}|user, domain)$ denote the estimated probability of a *user* having a *high*, *medium*, or *low* reputation respectively in category c_i given his/her *domain*. Similarly, the real values of these probabilities are represented by the following vector:

$$\mathcal{R} :< P(C_{i,h}|user, domain), P(C_{i,m}|user, domain), P(C_{i,l}|user, domain) > . \tag{18}$$

These vectors are calculated using the global reputation Markov tree and the technique described in the previous subsection. The degree of similarity of these two vectors ($\hat{\mathcal{R}}$ and $\mathcal{R}$) is calculated using the Cosine similarity measure; therefore, all similarity values are in the range of [0, 1]. Let $AUC(user, domain)$ denote the area under the curve of the line chart of similarity versus categories

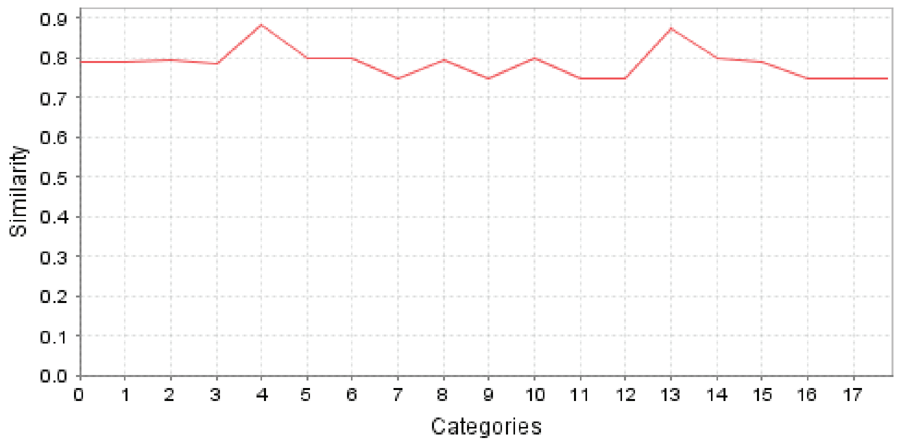

Fig. 6. Similarity between the real and approximated reputations for different categories.

(see Figure 6) of a specific *user* and his/her *domain*, and further let $\sigma(\hat{R}, R)$ be the Cosine similarity between the real and approximated reputation vectors, since $0 \leq \sigma(\hat{R}, R) \leq 1$; therefore, $max\,(AUC(user, domain)) = |Categories|$, where $Categories$ is the set of all categories in our experiment (19 categories of the epinions.com dataset).

Based on the analysis presented above, let us define $\mathbb{E}(\sigma(\hat{R}, R))$ as the expected similarity value between the real and approximated reputations.

$$\mathbb{E}(\sigma(\hat{R}, R)) = \frac{AUC(user, domain)}{|Categories|} \tag{19}$$

The higher the value of $\mathbb{E}(\sigma(\hat{R}, R))$ is, the better our algorithm is performing in predicting user reputation values. To analyze the predicted values for our algorithm, we have plotted the values of $\mathbb{E}(\sigma(\hat{R}, R))$ for all users, and domains. This plot is shown in Figure 7. It can be observed from this figure that almost all values of $\mathbb{E}(\sigma(\hat{R}, R))$ are above 0.8, which demonstrates the high accuracy of our proposed technique.

5.3 Average Similarity Value Per User

Our previous experiment showed that our technique is competitive when the users' domain is determined in advance. However, a new question arises for cases where the user's domain is unknown. We would like to see how accurately the proposed scheme is able to predict reputations for a given a user in such circumstances. In order to analyze this condition, given a specific user, we have averaged all the similarities between the real and approximated reputations for all domains and categories were the user was present. Figure 8 demonstrates

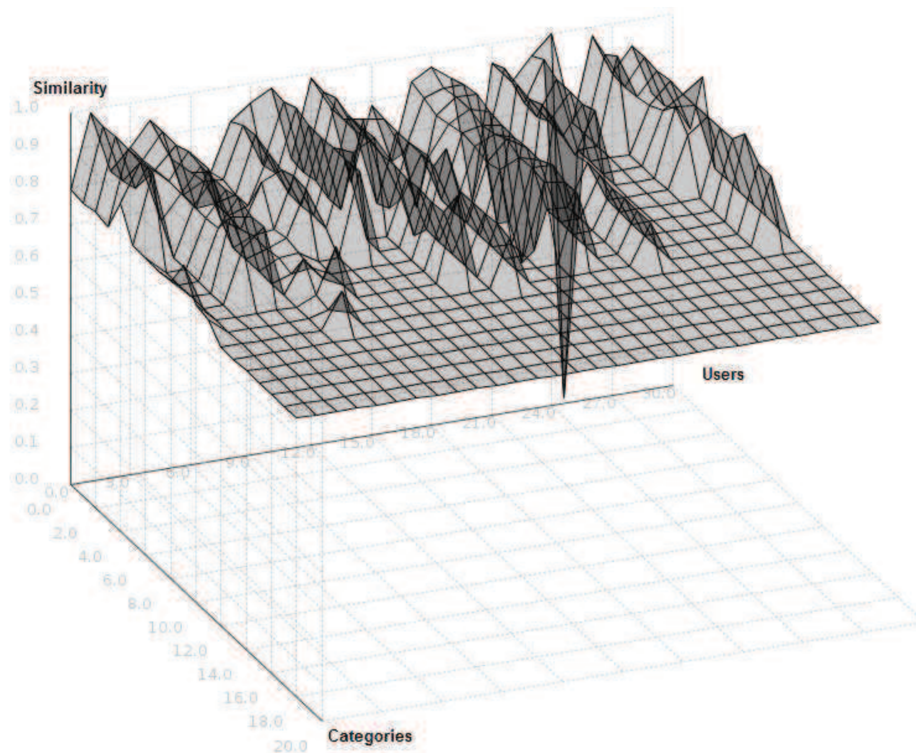

Fig. 7. Expected Similarity Values for Different Users and Domains

these average similarity values for a set of 30 random users. As it can be seen in this figure, in 93% of the cases (28 out of 30) the average similarity value is above 75%, which is quite reasonable when dealing with such a large dataset.

5.4 Average Similarity Value Per Category

Finally, how accurate are the reputation estimates when only the domain is known? To find an answer to this question, we have analyzed the average similarity values over a set of random users for different domains. For each domain, the average similarities between the real and approximated probabilities over all categories and users were calculated. These values along with their corresponding domains are shown in Figure 9. As it can be seen in this figure, no similarity value less than 0.8 can be found for any domain, which is quite promising in real world applications.

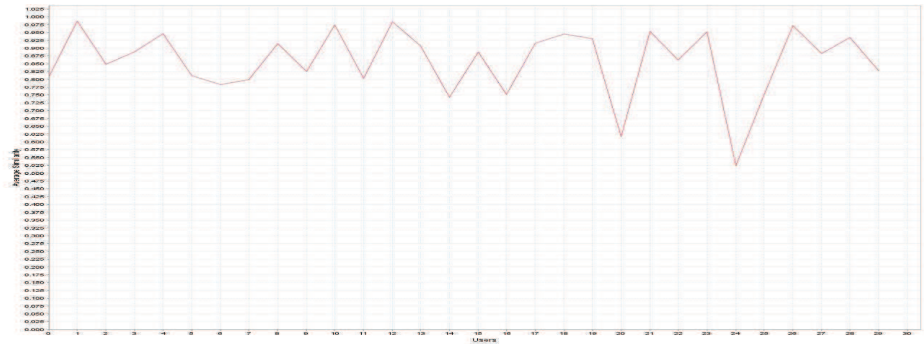

Fig. 8. Average Similarities per Users

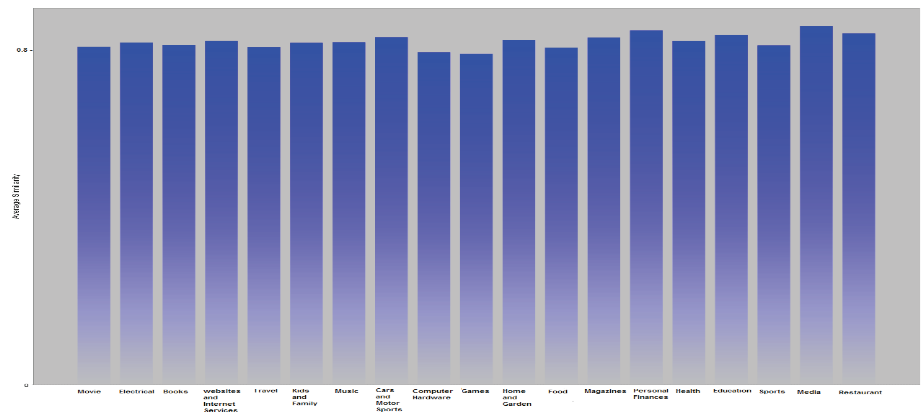

Fig. 9. Average Similarities per Categories

6 Concluding Remarks

In this paper, we have proposed a global reputation representation structure, and a reputation estimation model for multi-context environments. The proposed structure is based on the idea of valuation networks, and employs the belief propagation scheme introduced in the Shenoy-Shafer architecture [14]. The model is suitable for online communities that constitute multiple contexts. Currently, most such communities have only envisioned a single reputation value for each person. This has the disadvantage that users with a high reputation in an irrelevant issue can enter a new context where they do not possess any expertise and effect the community (possibly in a negative manner by deceiving the members of that context). In such cases, it is very hard for the members of the community to assess the reputation of the new member of their context by simply observing his previous reputation values. In a three-scale reputation formalization a lucky guess by the other members is only successful in 33.3% of the cases. Our model assists the users by estimating reputation values of new users in the contexts where their performance has not yet been observed. The proposed model has shown a 76.923% accuracy rate on a dataset collected from the epinions.com website.

As future work, we are interested in analyzing our proposed model with larger datasets to see how the model scales to larger systems. As shown in the figures depicting the structure of the epinions.com dataset, this dataset becomes rather parse as more data is extracted from the website; therefore, we will be looking at other larger website for evaluating the proposed model in the future. The experiments in this paper are preliminary, but show that the model can enhance the current state of the art in estimating reputation values. Furthermore, there are cases where an outsider brings more insight into the community of experts by pointing to an unknown concept for the community. We are interested to see how our model can be augmented to formally support such a situation.

References

1. Bagheri, E., Barouni-Ebrahimi, M., Zafarani, R., and Ghorbani, A. A. A belief-theoretic reputation estimation model for multi-context communities. In *Canadian Conference on AI* (2008), pp. 48–59.
2. Bagheri, E., and Ghorbani, A. A. Behavior analysis through reputation propagation in a multi-context environment. In *International Conference on Privacy, Security and Trust (PST'06)* (2006), ACM.
3. Bagheri, E., and Ghorbani, A. A. Exploiting trust and suspicion for real-time attack recognition in recommender applications. *iTrust/PST 2007* (2007).
4. Carter, J., Bitting, E., and Ghorbani, A. A. Reputation formalization for an information-sharing multi-agent system. *Computational Intelligence 18*, 4 (2002), 515–534.
5. Despotovic, Z., and Aberer, K. A probabilistic approach to predict peers' performance in P2P networks. *Lecture notes in computer science* (2004), 62–76.
6. Despotovic, Z., and Aberer, K. P2P reputation management: Probabilistic estimation vs. social networks. *Computer Networks 50*, 4 (2006), 485–500.
7. Diaconis, P. A mathematical theory of evidence. (glenn shafer). *Journal of the American Statistical Association 73*, 363 (1978), 677–678.
8. Kinateder, M., and Rothermel, K. Architecture and algorithms for a distributed reputation system. 2003, pp. 1071–1071.
9. Lehmann, N., and Haenni, R. An alternative to outward propagation for dempster-shafer belief functions. In *ECSQARU '95* (1999), pp. 256–267.
10. Massa, P., and Avesani, P. Controversial users demand local trust metrics: An experimental study on epinions.com community. In *AAAI* (2005), pp. 121–126.
11. Mui, L., Mohtashemi, M., and Halberstadt, A. A computational model of trust and reputation. *System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on* (Jan. 2002), 2431–2439.
12. Pearl, J. Reasoning with belief functions: An analysis of compatibility. *Int. J. Approx. Reasoning 4*, 5-6 (1990), 363–389.
13. Sabater, J., and Sierra, C. Review on computational trust and reputation models. *Artif. Intell. Rev. 24*, 1 (2005), 33–60.

14. Schmidt, T., and Shenoy, P. P. Some improvements to the shenoy-shafer and hugin architectures for computing marginals. *Art. Intel. 102*, 2 (1998), 323–333.
15. Seigneur, J.-M., Gray, A., and Jensen, C. D. Trust transfer: Encouraging self-recommendations without sybil attack. 2005, pp. 321–337.
16. Shenoy, P. P. Valuation network representation and solution of asymmetric decision problems. *Euro. J. Operation. Res. 121*, 3 (2000), 579–608.
17. Vu, L., and Aberer, K. A probabilistic framework for decentralized management of trust and quality. *Lecture Notes in Computer Science 4676* (2007), 328.
18. Wang, Y.; Vassileva, J. Trust and reputation model in peer-to-peer networks. *Peer-to-Peer Computing, 2003. (P2P 2003). Proceedings. Third International Conference on* (1-3 Sept. 2003), 150–157.
19. Yu, B., and Singh, M. P. Distributed reputation management for electronic commerce. *Computational Intelligence 18*, 4 (2002), 535–549.
20. Yu, B., and Singh, M. P. Detecting deception in reputation management. In *AAMAS '03: Proceedings of the second international joint conference on Autonomous agents and multiagent systems* (New York, NY, USA, 2003), ACM, pp. 73–80.

Table 1. The specification of the collected data from epinions.com

	Category	# Ratings	# Users	Rating Score				
				1	2	3	4	5
1	mvie	70641	11324	5423	7235	11471	21294	25218
2	kifm	49198	7814	3032	3556	4608	13172	24830
3	Music	43492	7641	1253	2140	4968	13488	21643
4	Books	43350	8538	1220	2054	4773	12935	22368
5	trvl	31467	8167	1480	1871	3417	9645	15054
6	game	26631	6397	1140	2037	3480	8427	11547
7	webs	23959	8406	3991	2992	3001	5924	8051
8	elec	23790	8599	1285	1564	2580	8272	10089
9	cmhd	18171	6673	1043	1242	1829	5790	8267
10	hmgd	16609	5586	1114	1028	1308	4411	8748
11	auto	15292	7309	644	864	1425	4870	7489
12	member	12228	4424	799	848	1262	3557	5762
13	well	10324	3089	690	850	1132	2966	4686
14	food-drink	7996	3527	714	818	1087	2462	2915
15	sport-outdoor	7459	2660	218	370	726	2417	3728
16	mags-news	6438	3524	369	578	681	1787	3023
17	finc	6181	3405	1175	651	611	1513	2231
18	rest	5678	2374	237	368	737	2036	2300
19	educ	4762	2712	249	436	642	1603	1832
20	pets	4415	1844	301	305	403	1144	2262
21	spirits	4235	1584	324	331	522	1286	1772
22	media	4020	2384	441	443	448	1123	1565
23	Cosmetics	3484	1075	248	349	419	1025	1443
24	Software	3252	2085	167	196	310	935	1644
25	health-aids	2942	1579	243	218	246	757	1478
26	sprt	2501	1302	88	192	325	832	1064
27	tools-access	1904	728	55	88	172	531	1058
28	health-beauty	1688	1332	182	156	217	460	673
29	btech	1651	1234	199	151	179	476	646
30	inst	1328	514	20	54	122	433	699
31	fddk	956	312	63	95	195	374	229
32	otdr	838	491	44	39	75	251	429
33	movies-theater	797	497	103	96	200	236	162
34	food&drink	638	368	60	50	84	222	222
35	kitchen-appliance	564	470	65	34	40	159	266
36	beat	562	327	45	57	92	150	218
37	Fragrances	530	306	9	20	36	120	345
38	bags	487	387	11	15	31	128	302
39	offc	406	210	13	23	43	116	211
40	Gifts	310	245	15	15	35	94	151
41	outdoor	283	146	3	15	21	64	180
42	Other Categories	1941	1715	123	132	175	547	964
	SUM	463,398	133,304	28,898	34,583	54,128	138,032	207,757