Towards a Global Online Reputation

Hui Li
University of Ottawa
55 Laurier Ave E
Ottawa, ON K1N 6N5 Canada
+1 (613) 562 5800, 8834
Hli103@uottawa.ca

Morad Benyoucef
University of Ottawa
55 Laurier Ave E
Ottawa, ON K1N 6N5 Canada
+1 (613) 562 5800, 4787
benyoucef@telfer.uottawa.ca

Gregor v. Bochmann
University of Ottawa
88 King Edward Ave
Ottawa, ON K1N 6N5 Canada
+1 (613) 562 5800, 6205
bochmann@site.uottawa.ca

ABSTRACT

Today's online reputation systems lack one important feature: globality. Users build a reputation within one community, and sometimes several reputations within several communities, but each reputation is only valid within the corresponding community. Moreover, such reputation is usually aggregated by the online platform's provider, giving the inquiring agent no say in the process. This paper proposes one way of dealing with this problem. We introduce an online reputation centralizer that collects raw reputation data about users from several online communities and allows for it to be aggregated according to the inquiring agent's requirements, using a stochastic trust model, and taking into account factors that qualify a user's reputation.

Categories and Subject Descriptors

C.2.0 [**Computer-Communication Networks**]: Security and Protection; H.3.5 [**Online Information Services**]: Web-based services.

Keywords

Reputation System, Online Trust, Stochastic Model.

1. INTRODUCTION

The Internet has enabled the proliferation of online interpersonal and business interactions between individuals who have never interacted before. These interactions are usually completed with some concern given that private information and the exchange of money and goods are involved. A mechanism is therefore needed to build trust among strangers who interact online. Trust can be divided into direct and recommender trust. While direct trust comes from direct experience, recommender trust is derived from word-of-mouth recommendations [1]. Trust is dynamic and can be developed over time as "the outcome of observations leading to the belief that the actions of another may be relied upon" [3].

One way to foster trust in online interactions is through collecting and managing information about the past behaviour of interacting parties. This information is then aggregated into an entity called reputation. Reputation is defined as a collective measure of trustworthiness based on the ratings of community members [2] which might affect the interacting party's future payoffs [4].

Online reputation systems are community tools that "collect, distribute, and aggregate feedback about participants' past behaviour" [5]. A negative reputation system gathers and distributes feedback on untrustworthy participants to discourage their behaviour; while a positive reputation system encourages participants with a history of honest behaviour [6]. In a hybrid reputation system, both positive and negative behaviours are taken into account. In such a case, participants start with neutral reputation values, then points are taken away as a punishment for bad behaviour or added as a reward for good behaviour [7]. EBay's feedback forum (www.ebay.com) is an example of a hybrid reputation system. It allows participants to rate each other with +1 for positive, 0 for neutral, and -1 for negative feedback. All the feedback values are then aggregated into one reputation value to be consulted by members of the eBay community [2].

Three entities are usually involved in trust models for online reputation systems: (1) the ***querying agent***, who is the user who wants to know whether a given user (the ratee) can be trusted; (2) the ***ratee***, who is rated by others on his/her past behaviour; and (3) ***the rater***, also called ***recommender***, who is the user providing information about the ratee, usually after having interacted with him/her.

Online reputation systems raise numerous challenging research questions [4]. In this paper we address one of them: the lack of globality. It is indeed difficult to exchange reputation data between different online reputation systems [5]. A member of the eBay community, for instance, cannot use his/her reputation outside the eBay community - hence the name "*local reputation*". It is desirable that a user who has a good reputation within one community could use his/her reputation within other communities - hence the name "*global reputation*".

As a step towards globality, we suggest the aggregation of reputation data from different online communities. A major difficulty is that each community calculates reputation differently. For instance, a rating value on eBay is between -1 and 1 while other online communities use ratings between 0 and 5 and may include textual comments as well. In order to aggregate reputation data from various communities, we propose a common reputation model into which the data can be translated.

Within a "global view" of online reputation, a ratee grants permission (with the possibility of opting out) to the communities where he/she has developed a reputation to share his data with a global aggregation service. We envision such a service to be offered by a third party who partners with online communities. The business, privacy, and security implications of an aggregation service are undoubtedly important but they are beyond the scope of this paper. Before interacting with the ratee, a user (the querying agent) logs into the aggregation service, looks up the ratee, gets access to his/her raw reputation data from partnering online communities, and configures the aggregation process. This process can be configured by parameters such as: weights assigned to online communities (perhaps giving more weight to the more established communities), transaction dates (perhaps giving more weight to the more recent recommendations), transaction values (perhaps giving more weight to recommendations for high value transactions), etc. Instead of providing a "dead" reputation score (as most of today's online reputation systems - e.g., eBay's feedback forum), the querying agent is given the opportunity to be involved by configuring the aggregation process and thus will find the aggregated feedback more useful.

The rest of the paper is organized as follows. Section 2 reviews discrete reputation models. Section 3 details our proposed reputation model. Section 4 reports on an implementation of the proposed model. Section 5 reviews related work and contrasts it with our proposed solution, and Section 6 concludes the paper.

2. DISCRETE REPUTATION MODELS

Computation models for online reputation can be classified into summation, weighted average, fuzzy, flow, Bayesian, belief, and discrete models [2]. We only discuss discrete models here.

Using a discrete model such as the one in [13], a rater evaluates his/her interaction with the ratee as "excellent", "good", "normal", "bad" or "worst". One shortcoming of discrete models is that they are not precise since "*heuristics mechanisms like lookup tables must be used*" [2] to convert feedback values into their numeric equivalent.

In [14] discrete feedback is used in conjunction with a statistical model to compute trust based on self-experience and recommendations from raters. It is assumed that the space of possible outcomes of transactions is finite (e.g., "excellent", "very good", "good" and "bad") and that N transactions have been observed for the same ratee by the querying agent or other raters. Assuming that ratee b will perform in a similar manner in the future, one can predict the probability of the different outcomes for future transactions using the formula:

$T_b(o)$ = (number of times the observed outcome was equal to o) / N.

$T_b(o)$ is the probability that a future transaction with ratee b will lead to an outcome o. The sum of the values $T_b(o)$ over all values of o yields the value one. $T_b(o)$ is also called the "trust that ratee b will provide an outcome o".

Instead of keeping all previous transaction outcomes in memory, an incremental trust update formula is used [14]. The current trust $T_b(o)$ (for each value of o) and the number of observations to date are kept in memory, and after a new transaction yielding outcome o is observed, the trust values and N will be updated as follows:

$T_b(o) = (T_b(o)*N + 1)/ (N+1)$.

$T_b(o') = (T_b(o')*N)/ (N+1)$ for o' different from o.

$N=N+1$.

Note that a multidimensional reputation model can be considered in the context of discrete reputation. For instance, a seller's reputation can be evaluated according to two dimensions: quality of good and service. For both dimensions, one may set up discrete values for the possible outcomes, such as "excellent", "good", etc.

3. PROPOSED REPUTATION MODEL

Our approach first aggregates a ratee's ***local*** reputations then combines them into a ***global*** one.

A ratee's ***local reputation*** is linked to a single community (e.g., a seller's reputation on eBay is considered local to the eBay community). If a community maintains an online reputation system, then the ratee is rated every time he/she transacts within that community. Note that we are not interested in the aggregated reputation value as provided by the community's reputation system but rather in the raw data. Let us assume that the raw data is comprised of the following elements.

- *Feedback value*: this is an essential parameter in reputation models (also called rating or recommendation). This value is typically given by a rater as feedback on a single transaction with the ratee. Reputation systems differ in their feedback representation formats, which could be discrete or continuous; numerical or textual or both. Some systems use feedback values alone to aggregate a user's reputation without considering other attributes (e.g., eBay only sums up the feedback values).

- *Information on rater credibility*: the quality of recommendations in trust systems is not guaranteed since nothing prevents malicious raters from providing unfair recommendations. As stated in [8, 9, 10, 16], feedback from raters with higher credibility should be weighted more than feedback from those with lower credibility since these are more likely to submit dishonest feedback. However determining rater credibility is a challenge. Shi et al. [15] for instance use data analysis and machine learning techniques to detect unfair recommendations. The querying agent may also compare the recommendation with his own experience. If the querying agent decides to interact with a ratee based on a recommendation from a rater, the difference between the rater's and the querying agent's perceptions, called semantic distance [1], can be used to adjust future recommendations from the same rater. In [9], raters' credibility is a function of their reputation within the community, hence reputable raters are considered more credible, and therefore their ratings weigh more.

- *Context factor*: various transaction parameters such as the size and time of a transaction can be considered, for instance the feedback for larger and more recent transactions may be assigned more weight. More recent transactions are likely to better reflect the current behaviour of the ratee [12, 16]. The size of the transaction [16] is considered in order to avoid the situation where a user behaves honestly for small transactions and dishonestly for larger ones.

- *Number of transactions*: the number of transactions is useful because the total feedback divided by the number of transactions reflects a ratee's reputation better than the total feedback alone.

It is important to note that other elements can be part of the raw reputation data hence it should not be limited to the elements mentioned above. Some reputation models, for instance, consider that the longer a rater has been part of a community, the more weight should be given to his/her feedback on other members. Others value the feedback of raters with the most transactions (regardless of how long they have been in the community). For more on this topic, the reader is referred to [4].

We assume discrete feedback is used. For instance eBay uses the discrete values "1", "0" and "-1' to stand for "Positive", "Neutral", and "Negative". Discrete feedback needs to be normalized, so normalizing the three eBay discrete values within the range [0, 1] would yield the numerical values 1, 0.5 and 0. Unfortunately, normalization could lead to unrealistic results. For instance, one ratee may have five "Positive" (1), and five "Negative" (0) transactions, while another may have ten "Neutral" (0.5) transactions. If every feedback is equally weighted, these two ratees would end up with the same reputation value (namely 0.5), which does not reflect the reality.

For that reason, we decided to follow a different approach inspired by Shi et al. [14]. In order to represent discrete reputation better, we propose a ***stochastic trust model*** based on the assumption that the ratee behaves like a stochastic process, and the reputation value represents the expectation that the ratee will act accordingly in the future (see Section 2). We calculate (Formulas 1, 2, 3) the estimated probability of each possible distinct outcome ("Positive", "Neutral", or "Negative") for the action of the ratee taking into account the different rating attributes introduced earlier. We then sum up these values together with the corresponding numerical value (representing that outcome) (Formula 4). The aggregated reputation of ratee *i* denoted by R_i is calculated using the following formulas:

$$P_i(o) = \sum_{k=1, f_{ik}=o}^{I(i)} \frac{W_{ik}}{\sum_{m=1}^{I(i)} W_{im}} \quad (1)$$

$$W_{ik} = CR_{ik} * CF_{ik} \quad (2)$$

$$CF_{ik} = a*T_{ik} + b*S_{ik} + c.\ a,b,c \in [0,1] \,\&\, a+b+c=1 \quad (3)$$

$$R_i = \sum_{o \in O} P_i(o) * NumVal(o) \quad (4)$$

Here: $P_i(o)$ = the estimated probability that ratee *i* will provide the outcome *o* in the future; *O* = the set of possible outcomes, such as "excellent", "good", "average", "bad", and "very bad"; *I(i)* = the total number of transactions; f_{ik} = ratee *i*'s feedback value for transaction *k;* W_{ik} = the aggregation weight for ratee *i's* feedback value for transaction *k;* CR_{ik} = the credibility of the rater who rated ratee *i* for transaction *k (*note that ratee *i* can be rated many times by the same rater, but we only consider the rater's reputation at the moment transaction *k* is performed); CF_{ik} = the context factor for ratee *i*'s feedback value for transaction *k;* T_{ik} = the time context factor for ratee *i*'s feedback value for transaction *k;* S_{ik} = the size context factor for ratee *i's* feedback value for transaction *k; NumVal(o)* = the numerical value corresponding to the outcome *o* (using a lookup table).

Table 1. Feedback values, their corresponding f_{ik} *and* aggregation weights

K	f_{ik}	W_{ik}
1	"Positive"	1
2	"Neutral"	0.5
3	"Negative"	0.5
4	"Positive"	1
5	"Neutral"	1
6	"Negative"	1
7	"Positive"	1
8	"Positive"	0.7
9	"Neutral"	1
10	"Positive"	0.3

For an illustration, consider the example of a ratee *i* within a community *X* who has been rated 10 times (i.e., I(i) = 10) possibly more than once by the same rater. Table 1 shows the 10 feedback values as well as their corresponding f_{ik}, and the aggregation weights W_{ik} for each feedback value. Table 2 shows the mapping of discrete values into numerical values.

Table 2. Lookup Table

Discrete	*Numerical*
"Positive"	1
"Neutral"	0.5
"Negative"	0

The estimated probability of ratee *i* being "Positive", "Neutral" or "Negative" in future transactions can be calculated as follows:

$$P_i(Positive) = \sum_{k=1, f_{ik}=Positive}^{10} \frac{W_{ik}}{\sum_{m=1}^{10} W_{im}} = \frac{1+1+1+0.7+0.3}{8} = \frac{1}{2}$$

$$P_i(Neutral) = \sum_{k=1, f_{ik}=Neutral}^{10} \frac{W_{ik}}{\sum_{m=1}^{10} W_{im}} = \frac{0.5+1+1}{8} = \frac{5}{16}$$

$$P_i(Negative) = \sum_{k=1, f_{ik}=Negative}^{10} \frac{W_{ik}}{\sum_{m=1}^{10} W_{im}} = \frac{0.5+1}{8} = \frac{3}{16}$$

The local reputation of ratee *i* within community *X* has a value of 0.65625 as estimated below.

$$R_i = \sum_{o \in O} P_i(o) * NumVal(o) = \frac{1}{2}*1 + \frac{5}{16}*0.5 + \frac{3}{16}*0 = \frac{21}{32} = 0.65625$$

In order to apply the computation model, the attributes that serve in the aggregation need to be normalized. Reputation systems maintained by different online communities use different formats to represent these attributes. Before aggregating them, it is necessary to normalize them into numerical values using mapping tables or conversion formulas as proposed in [11].

After the local reputations for every online community have been calculated they are aggregated into a ***global reputation***. The global reputation (GR_i) is calculated as follows:

$$GR_i = \sum_{j=1}^{I(j)} R_{ij} * \frac{W_j}{\sum_{m=1}^{I(j)} W_m}$$

Here: R_{ij} = local reputation for ratee i within community j; W_j = the aggregation weight for community j; $I(j)$ = the number of communities considered. Note that assigning a weight of zero to a community discards it from the global reputation aggregation. We note that we assume here that a ratee can be globally identified throughout all communities. However, the raters only need to be indentified within their community where their credibility is supposed to be known. The same rater may occur in different communities with different identifiers and different local credibilities.

4. IMPLEMENTATION

We implemented and tested the reputation model described above in the form of an Online Reputation Aggregation System (ORAS).

The system is composed of the following components: User Interface, Administrator Interface, Aggregation Module, Mapping Module and Lookup Tables (see Figure 1 for the architecture).

The *User Interface* can be used by querying agents to register, enter the identity of the ratee to be looked up, select the rating attributes, their weights, etc. Through the User Interface, the querying agent can select the configuration parameters for the aggregation process, such as the values of *a* (importance of time context factor) and *b* (the importance of the size context factor), and for each community j included in the aggregation, the weight W_j for the reputation in that community and the lookup table $NumVal_j$ containing the numerical values of the different outcomes considered in that community.

The Administrator Interface can be used to setup *Lookup Tables*, calculation algorithms, mapping schemes, conversion parameters, etc. The *Aggregation Module* implements the algorithms used to compute the local reputation for every community as well as the global reputation. The *Mapping Module* normalizes raw reputation data into a common format using Lookup Tables. Finally, participating *Online Communities* create and expose *Web Services* that give access to the raw *Reputation Data* of the ratees (and only those) who have granted them permission to do so.

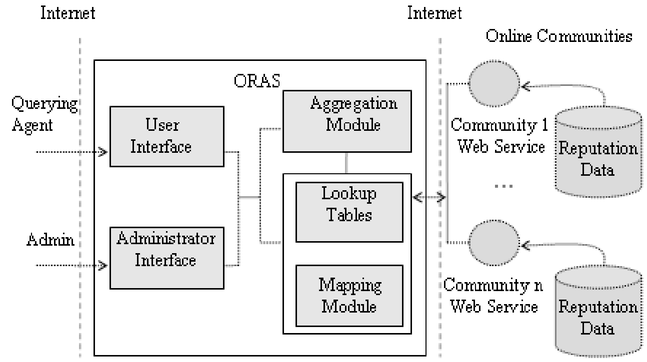

Figure 1. Conceptual Architecture of ORAS

Figure 2 shows how the user Hui@mail.com selects the communities (named X, Y, Z in this example) she wants to consider in her calculation. Remember that these communities are partnering with the aggregation service, and that the ratee in question (identified as Alex@mail.com in this example) has agreed to his data being shared with the aggregation service. In this example, the user assigns the highest weights to the communities believed to be the most accurate in reflecting the real reputation of the ratee.

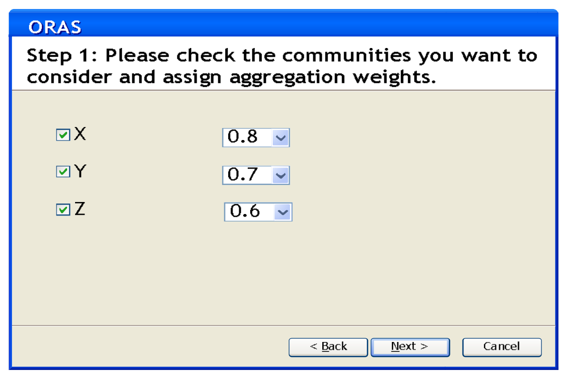

Figure 2. User Interface (Step 1)

In Step 2, the user chooses the rating attributes and sets their weights. In the current implementation of ORAS, the credibility of a rater is taken to be the value of his/her own reputation at the moment he/she provided the feedback. Other methods for computing rater credibility can be implemented.

Figure 3 shows the output screen after ORAS computes the local and global reputation values for ratee Alex@mail.com.

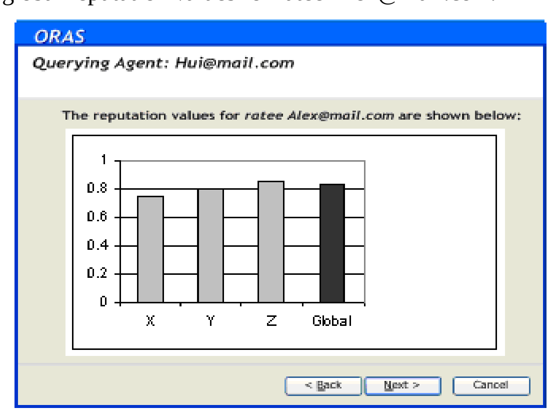

Figure 3. Local and global reputations are displayed

5. RELATED WORK

EgoSphere [17] is a reputation system aiming to integrate different reputation services by facilitating the transfer of reputation between them. It is composed of three modules: a web proxy, a reputation database and a reputation exchange. The web proxy runs on a user's computer monitoring all reputation-related activities. It fetches the webpage requested by the user from an EgoSphere-supported server, and analyzes the HTML code searching for reputation evidence and EgoSphere annotatable content such as usernames. The reputation database receives and manages the reputation evidence from many web proxy sources. The reputation exchange uses such evidence to calculate how much reputation data should be transferred from one service to another. The basic idea is that the more similarity two services have, the more reputation evidence can be transferred from one to the other.

Our solution is different in that the sharing of reputation information is conditioned by the user's approval, and our system does not need to parse HTML code because it has access to the raw data from participating online communities.

Commercial applications are being launched by online businesses (many of them start-ups) attempting to offer centralized reputation services, among them iKarma (http://www.ikarma.com) and authorati (http://www.authorati.com).

The Authorati rating service offers bloggers and online article authors a way to gain reputation and increase the visibility of their publishing. Users are allowed to list the URLs of their blogs/articles on their Authorati pages after registration. Readers can then rate the blogs/articles on Authorati. The rating consists of two parts: the authority rating (scale of 1-5) and the authorship rating (scale of 1-10). Each averaged rating will be shown below a blog or an article. Authorati allows readers to tag the contents of blogs/articles in fields such as arts, business, sports, technology, science, entertainment etc. In order to provide a portable rating service for blogs and online articles, Authorati offers its members a service for adding web widgets into their blogs or web pages to display the Authorati ratings. Members simply copy a piece of HTML code that generates the web widget and paste it on their blog, web page, or anywhere they want to show their Authorati ratings. Using a process that is more or less similar to Authorati, the iKarma online reputation service enables its members to rate other people and business. The idea is to provide a central location for managing reputation. In other words, when I interact with user U on website W, instead of rating him/her on website W, I go to a reputation centraliser (e.g., iKarma, Authorati) and enter my ratings there. Typically, I can also click on user U's badge/widget (if displayed on website W) to see his/her current reputation.

What we propose here is fundamentally different from what is currently offered by commercial services. Our solution (1) deals with raw reputation data; (2) offers the possibility to aggregate the local and global reputation according to the rater's specifications; (3) offers the possibility to select what communities (individual websites) to include in the aggregation process; and (4) provides a more configurable aggregation process for reputation.

The aim however remains the same: the portability, centralization, and globalization of online reputation.

6. CONCLUSION

This paper addressed the lack of globality in online reputation systems. Users who build a reputation in one community are unable to transfer it to another community. In view of the importance that reputation systems are gaining as a way of fostering trust in online business and interpersonal interactions, we believe globality to be an important feature. Our approach to achieve it is to gather raw reputation data about a ratee from various communities, aggregate the data from a given community into what we call a local reputation, then aggregate all local reputation values into a global reputation. The aggregation is based on options and weights which are selected by the inquiring agent according to his/her personal requirements. Our computation algorithm is based on a statistical model which takes into account several factors and parameters that qualify the reputation. A prototype based on the proposed model has been implemented and tested. The next step is to validate the model using real and/or simulated recommendation data.

Several extensions are envisaged for this work, among them: (1) considering reputation to be multidimensional where a ratee can be rated on more than one issue (product quality, service, etc.); (2) considering other factors in the aggregation of local reputation; and (3) investigating other ways to calculate the raters' credibility.

The novelty of our solution resides in the fact that it relies on raw reputation data from various online communities, relies on the ratee agreeing to (with the possibility of opting out) sharing his/her reputation data, involves the inquiring agent in the aggregation process for selecting various options, and uses a stochastic trust model in the aggregation process.

Finally, we note that several important aspects of global online reputation systems, such as business implications, privacy and security issues, and fraud prevention, are beyond the scope of this paper.

7. REFERENCES

[1] Abdul-Rahman, A., Hailes,S. Supporting Trust in Virtual Communities. In Proceedings of the 33rd Annual Hawaii International Conference on System Science. 2000. Hawaii.

[2] Josang, A., Ismail, R., Boyd, C., A Survey of Trust and Reputation Systems for Online Service Provision. Decision Support Systems, March 2007. 42(2): p. 618-644.

[3] Elofson, G., Developing Trust with Intelligent Agents: An Exploratory Study. In Proceedings of the first International Workshop on Trust, 1998: p. 125-139.

[4] Dellarocas, C., The Digitization of Word-of-Mouth: Promise and Challenges of Online Reputation Systems. Management Science, 2003. 49(10): p. 1407-1424.

[5] Resnick, P., Zeckhauser,R., Friedman, E. , Kuwabara, K. , Reputation Systems. Communications of the ACM, December 2000. 12: p. 45-48.

[6] Kollock, P., The Production of Trust in Online Markets. Advances in Group Processes, ed. E.J. Lawler, Macy, M., Thyne, S., and Walker, H.A. . Vol. 16. 1999: CT: JAI Press.

[7] Gilbert, A., Abraham, A., Paprzycki, M., A System for Ensuring Data Integrity in Grid Environments. Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'04), 2004. 1: p. 435-439.

[8] A.Malaga, R., Web-Based Reputation Management Systems: Problems and Suggested Solutions. Electronic Commerce Research, 2001: p. 403-417.

[9] Xiong, L., Liu, L., PeerTrust:Supporting Reputation-Based Trust for Peer-to-Peer Electronic Communities. IEEE Transaction on Knowledge and Data Engineering, July 2004. 16(7): p. 843-857.

[10] Zhou, R., Hwang, K., Power Trust: A Robust and Scalable Reputation System for Trusted Peer-to-Peer computing. IEEE Transactions on Parallel and Distributed Systems, June 21, 2006. 18: p. 460-473.

[11] Kamvar, S.D., Schlosser, M. T., Garcia-Molina, H., The Eigentrust algorithm for Reputation Management in P2P Networks Proceedings of the 12th International Conference on World Wide Web May 2003: p. 640-651.

[12] Change, E., Hussain, F. K., Dillon,T. , Reputation Ontology for Reputation Systems. OTM Workshops,LNCS 3762, 2005: p. 957-966.

[13] Manchala, D.W., Trust Metrics, Models and Protocols for Electronic Commerce Transactions. Processings of the 18th International Conference on Distributed Computing Systems, 1998: p. 312-321.

[14] Shi, J., Bochmann,G. v., Adams,C. , A Trust Model with Statistical Foundation. Processings of IFIP Workshop on Formal Aspects in Security and Trust, ed. M. T. Dimitrakos , F. Aug. 2004, Toulouse: Springer. 145-158.

[15] Shi, J., Bochmann,G. v. , Adams,C., Dealing with recommendations in a statistical trust model. Proceedings of AAMAS Workshop on Trust in Agent Societies, July, 2005: p. 144-155.

[16] Song, S., Hwang, K., Zhou, R., Kwok, Y., Trusted P2P Transactions with Fuzzy Reputation Aggregation. IEEE Computer Society, November, December 2005: p. 24-34.

[17] Bonawitz, K., Chandrasekhar, C., Viana, R., Portable Reputations with EgoSphere, MIT Internal Report, 2004.