

Evaluation of Trust in an eCommerce Multi-agent System using Fuzzy Reasoning

Andriy Hnativ and Simone A. Ludwig

Department of Computer Science, University of Saskatchewan, Saskatoon, Canada
{hnativ,ludwig}@cs.usask.ca

Abstract

Trust is a fundamental concern in large-scale open distributed systems such as multi-agent systems. It lies at the core of all interactions between the entities that have to operate in such uncertain and constantly changing environments. In this paper, an approach is developed for the evaluation of trust using fuzzy reasoning. The approach takes different trust sources into account, thereby minimizing the effect of wrong evaluations. It also incorporates a time factor for the evaluations of trust to address the different weighting of old versus new evaluations. Furthermore, the overall trust calculation consists of a non-linear weighted fuzzy calculation. A case study outlines different steps of the trust evaluation and shows, how the system computes the overall trust value, the reliability of the company, and the reliability of the evaluation.

1. Introduction

The impact of E-commerce trading is rising rapidly due to the enhancement of the internet and the customer's need to comfortably search and buy products online. Some of the advantages of e-commerce are: active pricing mechanisms, up-to-date databases, and streamlined procurement processes. An electronic market platform usually requires buyers and sellers to exchange offers-to-buy and offers-to-sell. The goal of this exchange is to reach an agreement on the suitability of a transaction between buyers and sellers. A transaction transfers one or more objects (e.g. a product, money, etc.) from one agent to another and vice versa [1]. However, conducting business via a computer platform brings in new challenges. One of the major shortcomings of electronic trade is that sales are being made without personal interaction. This means that consumers may buy goods from companies which they have not interacted before with, and whom

they do not know. Therefore, the platform needs to allow issues such as trust and reliability to be incorporated.

Trust is a major concern in large-scale open distributed systems. It is the central part of all interactions between the entities that have to operate in such uncertain and constantly changing environments. Both trust and reliability are imprecise and vague terms; therefore, people rate accounts of trust differently [2]. This leads naturally to the use of fuzzy logic as a description of these terms represent the human perception and reasoning process better. The main advantage of fuzzy logic is that it allows one value to be a member of more than one fuzzy set. Moreover, it allows values to have different degrees of membership in the fuzzy sets. In fuzzy logic, one can specify not only the terms, such as "cold" or "hot", but also modifiers such as "very cold" and "slightly hot". Therefore, our approach for the evaluation of trust and reliability in a multi-agent system for eCommerce is based on fuzzy logic.

We will be giving an account of related work in the area of trust assessment in multi-agent systems. Reece et al. in [3] developed a model to compute trust based on correlated multi-dimensional contracts. First, an agent estimates the utility of a contract. After that, the probability that the contract will be fulfilled is computed. This probability is calculated based on the Dirichlet distribution of agent's direct experience of the contract outcomes. Finally, the agent propagates the information inside a reputation system.

Robles et al. in [4] developed a trust model for a resource management multi-agent system designed for future mobile communication. A business model appropriate for selling of bandwidth resource and services is investigated. The model uses a notion of a concentric spheres structure. The core of this model contains the physical security. The middle layer incorporates the internal and the external security infrastructure. The outer spheres contain fairness,

reliability, reputation and loyalty to provide a complete model of trust for marketplaces.

Wiechers et al. in [5] presented an approach designed to create trusted infrastructures inside multi-agent systems. The idea is to organize a chaotic space into a hierarchical structure. There are special security distribution centers that are on top of this structure, which are elected by other agents. Also, the system has a pool of possible successors of those centers in case of a failure.

In [6] Huynh et al. proposed a trust reputation model in open multi-agent system that integrates some information sources to evaluate trust. It uses 4 kinds of trust: interaction trust, role-based trust, witness reputation, and certified reputation. This system assumes that agents provide information truthfully.

Maximilien and Singh in [7] state the problems of dynamic service selection and outline a multi-agent framework, where agents search for the services and select appropriate service providers based on user preferences and on trust levels. For this purpose qualities advertised by service providers are used.

The requirements of trust and reputation in the area of eCommerce have lead to our trust approach with the following features: (1) the approach takes different trust sources, with a finer granularity than other approaches into account, thereby minimizing the impact of wrong evaluations; (2) a time factor for the evaluations of trust is incorporated in our model to address the different weighting of old versus new evaluations and to allow old evaluation to expire; (3) the overall trust calculation consists of a non-linear weighted fuzzy calculation, evaluating the overall trust value and the reliability of both the company and the evaluation.

The remainder of this paper is structured as follows: Section 2 describes our fuzzy trust approach, outlining the components and the architecture. In Section 3, a case study is provided outlining how the trust evaluation is done using an example. Section 4 closes this paper by summarizing the findings and outlining future work.

2. Fuzzy Trust Approach

Our architecture consists of three components:

- trust categories;
- value of evaluation;
- time factor.

Based on these three components the architecture of this multi-agent system was designed and the evaluation of the overall trust computation is outlined.

2.1. Trust Components

2.1.1. Trust categories. For the trust evaluation we chose three trust categories: interaction trust, witness reputation, and certified reputation, such as given in [6]. However, we found that the witness category needs a finer granularity. A witness may be a collaborative company for a client, i.e. the company that is willing to provide truthful information to the client - in this case the client can trust this witness. A witness can be anonymous, meaning that the client does not know anything about the witness. In this case, the client may not trust this witness' evaluation as much as an evaluation of a collaborator, because it might be false (for example, this witness can be a partner for the evaluated company – in this case the evaluation may be too high, or a competitor for the evaluated company – in this case his evaluation may be too low).

There is also another case, whereby the provider of the evaluations may not be a collaborator of the client, but it is not anonymous either. In this case we know that this witness is not a competitor of the evaluated company or the evaluated company's collaborator. It could be a person that the client does not know and does not seem to be related to the evaluated company. This person may have some contact to allow the client to contact this person. But still the client does not have to trust this person the same way as this person trusts his/her partners. Therefore, we decided to divide our model further into the witness trust categories: trustworthy witnesses, not trustworthy but not anonymous witnesses, and anonymous witnesses.¹ All five trust categories are outlined in Figure 1.

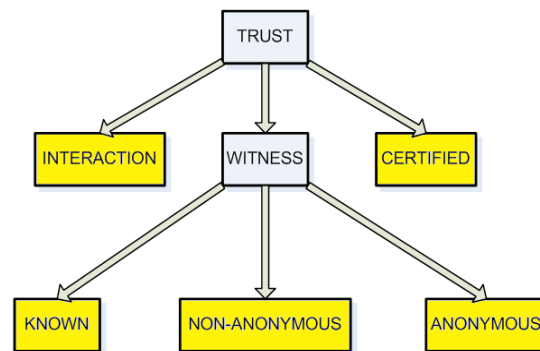


Figure 1. Five trust categories

¹ One case we want to mention in particular is the case where the evaluations are pseudo-anonymous, i.e., when there is just the name of the provider given, which does not give any reasonable amount of information. In this case, these values are put in the “anonymous witness” category.

2.1.2. Value of Evaluation. Each value of the evaluation describes the level of satisfaction a user has received from the service, which was provided by a particular company. This evaluation is given in linguistic terms. This approach is not typical, but we decided to use it, as the approach that involves numerical values has many shortcomings, as these values are very subjective. Even if two clients received exactly the same service they are likely to provide different numerical values for the trust, given the fuzzy and vague nature of our language². Moreover, even one company can provide different values at different times. On the other hand, if users describe their satisfaction in words, the descriptions reflect the true nature better³. Therefore, there will be significantly less ambiguity in our approach. This approach of using linguistic terms to describe knowledge leads us naturally to fuzzy logic. In our system, the evaluators describe the level of satisfaction using three terms “low”, “moderate”, and “high” which get mapped to three fuzzy sets accordingly. It may seem that this technique does not provide a good and precise assessment, or that it is just another variation of ebay’s evaluation system, whereby values +1 (positive), 0 (neutral), or -1 (negative) are used for the ranking [8]. But one advantage of the fuzzy approach is that each of these possible values can be weakened or strengthened by some additional terms, so-called *modifiers*. For example, the value “extremely high” is different than the value “somewhat high”, and this provides a more precise evaluation as the ebay system only allows a value of +1 to be assigned for both cases.

2.1.3. Time factor. Usually, but not always, more recent evaluations should matter more than older ones when aggregating all trust values from different sources. To include a time factor, using a linear or non-linear function to determine the importance of the input based on how long ago it was provided seems a decent decision, but there are also many drawbacks to this approach. We believe that there should not be a difference between the weightings of the current and the previous day’s evaluations for the overall trust calculation, because in this case any company can cheat the evaluation, providing very good feedback about itself each new day via certified trust and anonymous witnesses, and these values will count

² Assume that two clients are very satisfied with the service – what value will they provide? One company may provide the highest possible, while others will not, thinking that the highest values are only for exceptional cases.

³ Using the example in footnote 2, users do not have to provide a value for the service obtained, they can just rate the service provided with “very satisfied”.

much more than the true but older values. Therefore, we decided to create two sets of values based on the time when the evaluation was produced. One set represents more recent values, another set contains older values. The values from the first set count more than the values from the second set for the overall trust calculation. By default, the evaluations for the last 180 days, which is approximately half of a year, belong to the set of recent values; all the other evaluations belong to the set of older values. Furthermore, by default, more recent values count twice as much as older ones (the proportion between the weightings of older values and the weighting of more recent values is 0.5).

We realize that different applications and different services require different approaches. We believe, for particular services, which undergo frequent quality changes, such as internet provisioning or mobile connections, the value of 180 days is too high, while for others, for example time-consuming services such as stadium construction etc. it will be too small. The same is true for the proportion between the sets – 0.5 does not always provide the best representation. Thus, our system allows users to change both the number of the “more important days” and the proportion value. Moreover, it allows one to assign different values for the different trust categories.

2.2. Trust Architecture

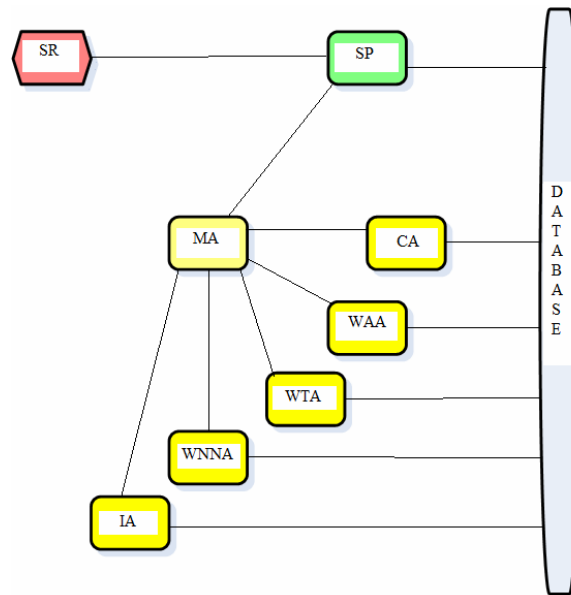


Figure 2. Trust architecture

Figure 2 shows the trust architecture within the multi-agent system. The interactions are as follows. The service requester agent (SR) asks the service

provider agent (SP) to get information about a particular service. The SP then checks in the database if this service is available. If it is not, the SP sends a message to the SR advising that the information regarding the requested service cannot be provided. On the other hand, if the service is present in the database, the SR receives the list of names of all the companies that provide the requested service from the database. Then this agent sends these company names one by one to the main agent (MA) which contains the fuzzy component. The MA upon receiving the name of the company sends it to the 5 fuzzy agents – IA (interaction trust agent), WNNA (witness not trustworthy not anonymous trust agent), WTA (witness trustworthy trust agent), WAA (witness anonymous trust agent), and CA (certified trust agent).

These 5 agents are responsible for the evaluation of the 5 different types of trust. All 5 agents are fuzzy agents that evaluate the trust of the company for a particular category and send the calculated trust value back to the MA. Then the MA combines the received overall trust value, computes the remaining values (the overall reliability of the company, linguistic description of the reliability, reliability of the evaluation), and sends all these values to the SP. The SP then creates the representation of the results (builds graphs and tables) and sends all this information to the SR. The sequence of actions is shown in Figure 3.

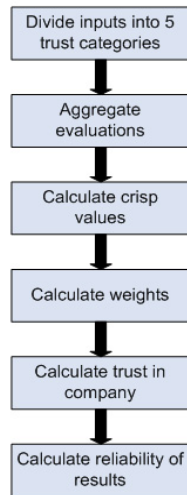


Figure 3. Trust calculation steps

All the categories are processed in the same manner. If an evaluation, which is a linguistic term, belongs to the set of recent values, then it is added to the overall trust fuzzy variable without any changes. If it belongs to the set of older values, then firstly the value is decreased according to the selected proportion, and is only afterwards added to the overall trust value.

Thus, in this way more recent values will have a greater impact than older ones.

Afterwards, we get five fuzzy variables that will represent five sets of trust. In order to combine them and to calculate the total value of trust, we need to defuzzify these values (to convert the fuzzy value to the crisp value). The crisp values that we get will be in the range from -1 to 1, where 1 means “absolutely reliable” and -1 means “absolutely unreliable”. A value can both increase the trust if it is positive (thus increasing the trust of the evaluated company), and decrease it, if it is negative (thus decreasing the trust of the evaluated company). The final value of the overall trust after normalization will be in the same range (-1;1).

When we aggregate the results of the different trust categories, these results must have different weights and therefore, we introduced coefficients for these five categories. First of all, our weights can change for different categories based on the number of evaluations. For example, if there is the same number of evaluations in the “own” and “trusted witness” categories, the “own” category weight must be higher than the “trusted witness” weight. But if there is only 1 evaluation in the “own” category and, say, 10 evaluations in the “trusted witness” category, then, we believe, the weight of the latter must be higher. We can trust the evaluations from the categories “own” and “trustworthy witness” more than the evaluations from the other categories. Thus, the weights of only these two categories are able to reach the maximum value for a coefficient of 1. The maximum possible values of the other three weights are lower; the lowest is the maximum value for the certified reputation category “evaluated person”, as the results of the evaluations from this category are usually biased. If there are no interactions in one category, the weight for this category will be equal to zero. If there are more interactions than a predefined value, then the weight will be equal to the maximum for that category. We stopped increasing the weight at this point, because it may be possible that other categories will not play an important role if the weight of one category is very large. For example, in the case of the “evaluated person” category, an evaluated person may give a large number of positive evaluations (many of which might be false) and then this category will determine the final trust in the company, which is not desirable. If the number of evaluations is between zero and a predefined number, then the weight will be determined by an equation, which is different for each category. Table 1 provides all equations for the different trust categories.

Category	Own	Evaluated Person	Trustworthy witness	Anonymous witness	Unknown non-anonymous witness
	$x = 8$	$x = 12$	$x = 36$	$x = 27$	$x = 8$
for $n = 0$	$w = 0$	$w = 0$	$w = 0$	$w = 0$	$w = 0$
for $0 < n < x$	$w = \frac{\sqrt[3]{n}}{2}$	$w = \frac{\sqrt[3]{n}}{4} - 0.2$	$w = \frac{\sqrt[3]{n}}{2} - 0.15$	$w = \frac{\sqrt[3]{n}}{12}$	$w = \frac{\sqrt[3]{n}}{3} - 0.2$
for $n \geq x$	$w = 1$	$w = 0.3$	$w = 1$	$w = 0.5$	$w = 0.8$

Table 1. Weight equations for all five trust sources

We determined the equations by the following reasoning. For our evaluation the law of diminishing returns holds – each new evaluation of the trust for a particular company bears a smaller return for the calculation of the overall trust than the previous one. Imagine two cases: (1) We have 0 evaluations for the company; (2) We have 500 evaluations. In the first case, the next evaluation will have a crucial impact on the final overall trust calculation, while in the second case the next evaluation will have only a minimal influence. Therefore, using linear functions is not the best choice because of the shown behavior. The best choice is to use root exponential functions - functions, where for every new value of x the value of y grows less than for the previous one.

Looking at the category “own” in Table 1, if there are no evaluations, the value of this coefficient will be equal to zero. We assume that 8 interactions is enough to understand the reliability of a company from direct interactions, so starting from eight evaluations in this category the weight will be equal to the maximum which is 1. The equation we chose for this weight, when the number of evaluations is between 1 and 8, is a cubic root function, as the behavior of this function allows us to reach the behavior we desire (the behavior of the law of diminishing returns). Thus, if there is 1 interaction, the weight will be equal to 0.5, which is half of the possible value, but still larger than the maximum weight of the certified trust. With every further evaluation there will be less growth of the weight.

The same reasoning is applied to the weights of the four other categories. We determined the sufficient number of evaluations for each category and when the actual number is larger than the predefined, the weight will take its maximum value. Again, we determined these values based on the simple reasoning: the less reliable the category is, the more interactions will be necessary and the smaller the weight value will be.

Now, having five crisp values for the five categories c_i , and the five corresponding weights w_i , the system calculates the final trust value, the

reliability of the company, using the following equation:

$$\text{overallTrust} = \sum_{i=1}^5 c_i \cdot w_i \quad (1)$$

After that, this value is mapped to the fuzzy sets of the level of reliability, and the corresponding linguistic variable is determined.

The system also calculates the reliability of the evaluation, using the following reasoning (simplified here for explanation): The fewer evaluations for the company were given, the less reliable are the evaluations. For example, if for some company only one evaluation is available in the database, the overall computed reliability of the company will be equal to that value, which of course might not be the true value. In this case, the system reports, that the reliability of the evaluation is “very low”.

As was discussed earlier, the values of different categories have different importance. For example, one evaluation in the interaction trust category is more important than two evaluations in the certified trust category. Taking this into consideration, the equation for the reliability of the evaluation looks as follows:

$$\text{reliabilityOfEvaluation} = \sum_{j=1}^5 n_j \cdot v_j \quad (2)$$

where n_j represents the number of evaluations in each of the five categories, and v_j are the predefined coefficients for the corresponding categories.

3. Case Study

Let us look at a case study to see how the system works. Assume that a user requests a stock brokerage service and that he uses standard settings (the number of days in the recent set will be 180, and the evaluations with an older timestamp will weigh half than the recent ones). After requesting this service, the user receives the following list of companies providing a stock brokerage service: Company1, Company2, Company3, Company4, Company5, Company6,

Company7, Company8. Then the system evaluates these companies one by one. Let us look at how the system evaluates “Company1”, given the information in the database shown in Table 2.

Evaluation	Provider	Trustworthy	Anonymous	Days ago
very high	certified	no	no	45
very low	certified	no	no	279
very high	witness	yes	no	945
moderate	witness	no	no	23
slightly high	witness	no	no	91
very high	witness	yes	no	2
extremely high	witness	no	no	38
very high	witness	yes	no	66
extremely high	witness	no	yes	94
very high	witness	yes	no	2
very high	witness	no	yes	1

Table 2. Trust information in database

For this company, the evaluations of 4 types exist: certified, trustworthy witness, anonymous witness, and not trustworthy and not anonymous witness. There are no type “own” evaluations in the database, i.e., the client did not have direct interactions with this company before, therefore, this category does not count for the final trust calculations, so its weight will be equal to 0.

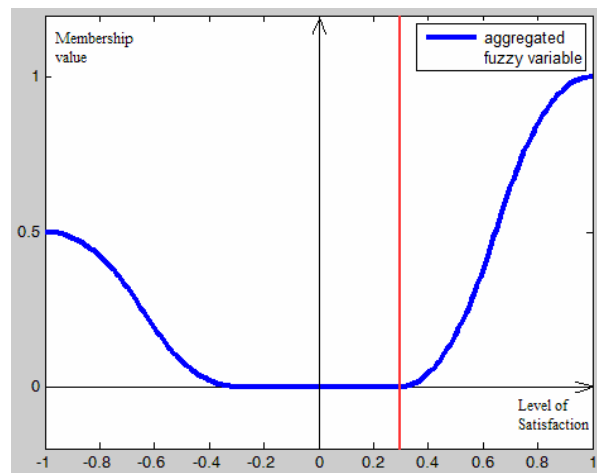


Figure 4. Fuzzy sets of level of satisfaction

For other categories the corresponding agents find evaluations and weights for these categories. For

example, for the certified reputation the aggregated fuzzy variable will look as shown in Figure 4.

The red colored vertical line shows where the defuzzified crisp value for this fuzzy variable fits in the graph, which is equal to 0.279. As it can be seen, it is closer to the high set than to the low set. There are only two evaluations in the “certified” category – “very high” (right hand side) and “very low” (left hand side). During the construction of the resulting fuzzy value, the value of the second evaluation counts half than the value of the first one, due to the fact that the first evaluation belongs to the category of recent values (since it was provided less than 180 days ago), and the second one belongs to the set of old values.

The weight of this category will be equal to $\frac{\sqrt[3]{2}}{4} - 0.2 = 0.115$ according to the equation as shown in Table 1 (third column).

The same process is applied to the other three categories. At the end, the system normalizes the weights, and calculates the final trust value according to Equation 1: $overallTrust = 0.705936$. The corresponding linguistic variable is “pretty high”.

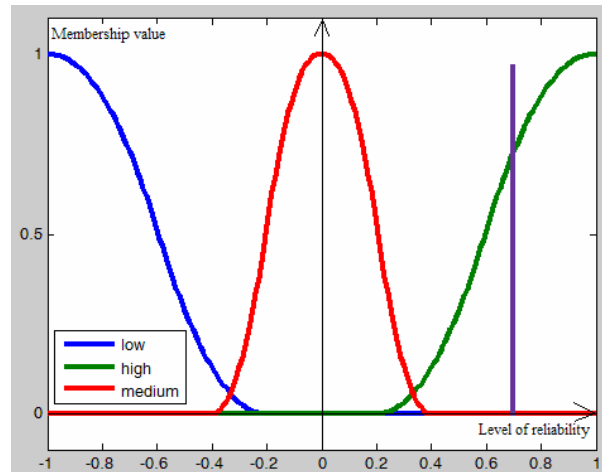


Figure 5. Fuzzy sets of level of reliability

The system computes the reliability of the evaluation according to Equation 2: $reliabilityOfEvaluation = 0.442$.

This value means that the reliability of the evaluation is pretty low, thus, our trust calculation was “not very reliable”, because of the fact that not enough evaluations were given for this company. The same procedure is applied to the other companies, in the end resulting in Table 3. As you can see, Company7 or Company3 seem to be the best choice for the user. They both have high values for reliability of the company and the evaluation. Both companies score

“extremely high” for the reliability of the company, even though, the overall trust value is slightly higher for Company7. In the “reliability of evaluation” category, Company3 scores higher with “very reliable” as opposed to “quite reliable” for Company7. This means that the user can be surer of the accuracy of the trust calculation for Company3, than for Company7, as evidently Company3 had more evaluations than Company7.

Name of company	Overall trust value	Reliability of company	Reliability of evaluation
Company7	0.851847	extremely high	quite reliable
Company3	0.847798	extremely high	very reliable
Company1	0.705939	pretty high	not very reliable
Company8	0.195775	moderate	not very reliable
Company9	0.168095	moderate	not reliable
Company5	0.102949	moderate	quite reliable
Company2	0.000000	moderate	not reliable
Company6	-0.466371	pretty low	quite reliable
Company4	-0.682032	low	not very reliable

Table 3. Results of trust calculation

4. Conclusion

Trust is a fundamental concern in a multi-agent system. This paper presented an approach for the evaluation of trust in such system using fuzzy reasoning. This approach incorporates many unique components. First of all, it uses fuzzy reasoning, which allows one to deal with vague concepts such as trust in a natural way. Secondly, it uses a finer granularity of trust sources, which was not used before, thus allowing to process data from distinct sources differently. Thirdly, the approach uses non-linear dynamic weight calculations, thus, the weights for the input categories for each company will be different. It uses an adjustable time factor, which makes this approach customizable. Finally, the system computes not only the trust of companies, but also the reliability of the evaluation, thus, a user receives valuable information before making a decision.

One limitation of our approach is our assumption that a service is evaluated based on the company which is providing this service. However, this is not always the case. One company might be specialized in one particular service, but may also offer other services which it is not specialized in. In this case, the particular service the company is specialized in will be trusted more than the other services of the company. Therefore, a finer granularity of the approach on the service level of companies has to be introduced.

5. References

- [1] M. Ströbel and M. Stolze, “A Matchmaking Component for the Discovery of Agreement and Negotiation Spaces in Electronic Markets”, *Journal of Group Decision and Negotiation*, 11:165-181, 2002.
- [2] F. K. Hussain, E. Chang, T. S. Dillon, “Trust Relationships and Reputation Relationships for Service Oriented Environments”, *Proceedings of IEEE International Conference on Computer Systems and Applications*, March 2006.
- [3] S. Reece, A. Rogers, S. Roberts, N. R. Jennings, “Rumours and Reputation: Evaluating Multi-Dimensional Trust within a Decentralised Reputation System”, *Proceedings of the 6th International Conference on Autonomous Agents and Multiagent Systems*, Honolulu, Hawaii, USA, May 2007.
- [4] S. Robles, J. Borrell, J. Bigham, L. Tokarchuk, L. Cuthbert, “Design of a trust model for a secure multi-agent marketplace”, *Proceedings of International Conference on Autonomous Agents and Multiagent Systems*, Montreal, Canada, 2001.
- [5] W. K. Wiechers, S. Daskapan, W. G. Vree, “Simulating the establishment of trust infrastructures in multi-agent systems”, *Proceedings of the 6th international conference on Electronic commerce*, Delft, The Netherlands, 2004.
- [6] T. D. Huynh, N. R. Jennings, N. Shadbolt, “Developing an integrated trust and reputation model for open multi-agent systems”, *Proceedings of the 7th International Workshop on Trust in Agent Societies*, (pp. 62-77), New York, USA, 2004.
- [7] E. M. Maximilien, M. P. Singh, “Agent-based trust model involving multiple qualities”, *Proceedings of International Conference on Autonomous Agents and Multiagent Systems*, Utrecht, The Netherlands, July 2005.
- [8] B. Rietjens, “Trust and reputation on eBay: Towards a legal framework for feedback intermediaries”, *Journal of Information and Communications Technology Law*, vol. 3, no. 1, pp. 55-78, 2006.