Journal of Theoretical and Applied Information Technology
15th November 2012. Vol. 45 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DESIGN AND IMPLEMENTATION OF P2P TRUST MODEL BASED ON THE CLASSIC BAYESIAN NETWORKS THEORY

[1]XIAOHONG SHI, [2] YANYI ZHANG

[1] Jiangxi Tourism & Commerce Vocational College, Nanchang 330100, Jiangxi, China

[2] Jiangxi Chang-Da High-Tech Information Industry Co., Ltd, Nanchang 330096, Jiangxi, China

ABSTRACT

Based on Bayesian network theorem, the paper proposed the novel trust model of P2P network named Trust-BT. The novel new Trust-BT model is based on the P2P network nodes' history of all types of transactions, prior experience, the use of Bayesian statistical analysis methods calculate the global trust value of every network node, select the node with high trust value node transactions. The mathematical analysis and the simulation results are shown that the model can effectively resist all kinds of malicious nodes attack, compared with the classic trust model (i.e. Eigentrust) improve the successful transaction rate of P2P network.

Keywords: *Transaction History, Bayesian Theorem, Posterior Probability Density, Hypothesis Testing*

1. INTRODUCTION

Peer-to-peer network technology widely used because of its open, anonymous, autonomy, security issues, resulting free-rider [1] a serious decline in the system's quality of service (QoS). For example, some of the nodes use of P2P the network diffusion viruses with forged documents; 70% of users never share any files, 50% rely on 1% share user query response file. An effective way to solve the above problem is a confidence-building mechanism, nodes with high trust value to obtain quality QoS [2]. The proposed Eigentrust and literature [3] proposed Peertrust is a classic trust model.

Eigentrust in the node of the global trust value calculated by iterative method for solving: $t_i^{(k+1)} = (1-\alpha)(c_{1i}t_1^{(k)} + c_{2i}t_2^{(k)} + ... + c_{ni}t_n^{(k)}) + ap_i$, c_{ji} projected the local trust value of node j to the node i, all transaction data between the node i and the node j, p_i said the initial trust value of the node i for the allocation factor α , but in Eigentrust in global trust value only a relative value, can not accurately show the degree of credibility of this node. Literature [4-5] in the basis of Eigentrust with Peertrust local improvements, but did not overcome the inherent defects of the two algorithms, is still using the iterative method to calculate the global trust value by the local trust value, each transaction should cause the entire network iterative calculation of the node, computational complexity, traffic, and faced with the transaction data is too sparse, the calculation is not precise enough and easy prey for malicious node impersonation, slander and conspiracy attacks.

The network is huge so not prone to repeated transactions between any two nodes, more feedback on a node, the easier wrong bias evaluation expelled. The aggregation node all nodes directly experience obtained global trust value? This paper is based on Bayesian theory [6-7] proposed a new P2P global the trust model Trust-BT (Trust Model based on Bayesian Theory), the model overcomes the computational complexity of the trust model based on the iterative method, the global trust value expressly node trustworthy and comprehensive consideration of the malicious node attacks and other security issues.

Bayesian statistical theory to overcome the classical statistics requires a large sample rather sparse sample of the problems in the real environment, it's the basic idea: the object overall sample have a certain understanding with the a priori probability, based on the current samples of correction for the a priori probability, the posterior probability distribution, and a variety of statistical inference on the basis of the posterior probability.

The probability density function $p(x|\theta)$ is set target overall priori probability density function $\pi(\theta)$ of the parameter θ. The observed value of the sample X is obtained with $x = (x_1, x_2 ... x_n)$, the sample and parameters of the joint probability density distribution $h(x,\theta) = \prod_{i=1}^{n} p(x_i|\theta) \times \pi(\theta)$ as a

Journal of Theoretical and Applied Information Technology
15th November 2012. Vol. 45 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

marginal probability density distribution $m(x)=\int_{\Theta} h(x,\theta)d\theta$ of the sample, then the θ posterior probability density distribution:

$$\pi(\theta\left|x\right.)=\frac{h(x,\theta)}{m(x)}=\frac{p(x\left|\theta\right.)\pi(\theta)}{\int_{\Theta}p(x\left|\theta\right.)\pi(\theta)}$$

This is in the form of the probability density function of the Bayesian formula.

In-depth study and application of trust model based on Bayesian theory in wireless sensor networks and Ad-hoc network [8-9], and achieved good results. Wireless sensor networks, each node in the P2P network transaction behavior, not only with their neighbors with any node in the network may be a peer-to-peer communication, increase the P2P trust modeling based on Bayesian theory complexity.

2. TRUST MODEL BASED ON BAYESIAN THEORY (TRUST-BT)

2.1 The Calculation of Trust Value

Define 1: nodes of the global trust value

In the current time period $T(t)$, according to the probability of transaction historical evaluation of all nodes with the node in the network egress node is willing to provide services, known as the node i's global trust value θ_i, $0\le\theta_i\le 1$, any node has a globally unique trust value.

Defined 2: trading record

Two nodes in the network each occurrence of a transaction behavior, the party receiving services will give the party providing services do behavior record node i to node j behavior recording $x_{ij}=\{s_{ij},f_{ij}\}$, wherein s_{ij} represents the node j to the service node i the number of successes, f_{ij} node j to node i service the number of failures.

Define 3: node i transaction number of samples $n=\sum_{j=1}^{N}(s_{ji}+f_{ji})$, where N represents the total number of nodes in the network.

Brought together a time period $T(t)$, all nodes in the network recorded the behavior $x_{ji}(j=1,2..n)$ of node i, estimated the node i in the current global trust value $\theta_i(0<\theta<1)$.

Any node i to provide services to trade $X_k=\begin{cases}1, success\\ 0, failure\end{cases}$ $k=1..n$, then the probability distribution of each transaction X_k:

$$p(X\left|\theta_i\right.)=\begin{cases}\theta_i^{x}(1-\theta_i)^{1-x} & x=0,\ 1\\ 0 & otherwise\end{cases}$$

Let $y=\sum_{j=1}^{n}s_{ij}$, then the joint probability function of $X_1,...X_n$:

$$p_n(X\left|\theta_i\right.)=\theta_i^{y}(1-\theta_i)^{n-y}$$

Saurabh [10] and Josang [11] and confirmed *beta* density function is suitable for its simplicity, flexibility, and strong statistical theory foundation to build a stable and reliable trust system.

Taken θ_i priori probability density function $p(\theta_i)$ obey *beta* distribution $Beta(a,b)$:

$$p(\theta_i)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta_i^{a-1}(1-\theta_i)^{b-1}$$
$$\forall 0\le\theta_i\le 1, a\ge 0, b\ge 0$$

θ_i is the posterior probability density function:

$$p_n(\theta_i\left|X\right.)=\frac{p_n(X\left|\theta_i\right.)p(\theta_i)}{\int_0^1 p_n(X\left|\theta_i\right.)p(\theta_i)d\theta_i}$$
$$\propto p(X\left|\theta_i\right.)P(\theta_i)\propto\theta_i^{a+y-1}(1-\theta_i)^{b+n-y-1}$$

θ_i of the posterior probability density function $p_n(\theta_i\left|X\right.)$ subject to parameters for the *beta* distribution of $a+y, b+n-y$.

θ_i Bayesian estimation is the mean of the subsequent posterior distribution of $p(\theta_i\left|X\right.)$.

$$E(\theta_i)=\int_0^1\theta_i p(\theta_i\left|X\right.)d\theta_i$$

θ_i is the posteriori average:

$$E[\theta_i]=\frac{a+y}{a+b+n}$$

I.e. trust value of w in the time period $T(t)$ Bayesian estimation value $\frac{a+y}{a+b+n}$. If θ_i priori distributions [0,1], a uniform distribution, i.e. $a=b=1$ then priori mean of 0.5, i.e. the

Journal of Theoretical and Applied Information Technology
15th November 2012. Vol. 45 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

knowledge of nodes probability identified services it provides a probability of 0.5, which give each new node of the same opportunity to provide services or receive services.

2.2 The Examples Analysis

The trust value θ_i of the estimated node i obtained within this period of time $T(t)$, the transaction records are as follows:

$$x_{1i}=\{3,7\}, x_{2i}=\{3,10\}, x_{3i}=\{2,5\},$$
$$x_{4i}=\{4,14\}, x_{5i}=\{6,1\}, x_{6i}=\{5,11\}$$
$$x_{7i}=\{1,2\}, x_{8i}=\{5,16\}, x_{9i}=\{2,6\},$$
$$x_{10i}=\{6,17\}, x_{11i}=\{7,20\}, x_{12i}=\{3,10\},$$
$$x_{13i}=\{1,4\}, x_{14i}=\{2,7\}, x_{15i}=\{3,9\},$$
$$x_{16i}=\{8,9\}, x_{17i}=\{9,27\}$$

Supposing θ_i priori distribution of conjugated priori $Beta(a,b)$ wherein $a=10, b=12$, and θ_i is the posterior mean estimate.

Number of successful transactions in transaction records x represents the total number of transactions represented by n, the posterior distribution of x line with the binomial distribution $b(n,\theta_i)$, then θ_i:

$$\pi(\theta_i \mid x)=\frac{\Gamma(n+a+b)}{\Gamma(x+a)\Gamma(b-x+n)}\theta^{x+a-1}(1-\theta)^{n-x+b-1}$$

That $\pi(\theta_i \mid x) \sim Beta(x+a, n-x+b)$, $\frac{x+a}{n+a+b}$ is estimated to be easy to know the mean of θ_i.

The calculated

$$x=3+3+2+4+6+5+1+5+2+6$$
$$+7+7+3+1+2+3+8+9=77$$
$$n=10+13+7+18+7+16+3+21+8+23$$
$$+27+13+5+9+12+17+36=245$$

Data into the draw θ_i posterior mean is estimated to be 0.3258.

3. TRUST-BT SECURITY

P2P networks, self-organization, Anonymous, the characteristics of a high degree of turbulence (churn) Trust-BT there are security issues:

(1) Free-riding attacks: selfish nodes unrestricted access to services from selfless the node or mixed node, but never provide a service to the system.

(2) Slander attack: malicious nodes deliberately transaction with node negative evaluation, such as the success of the transaction is recorded as unsuccessful.

(3) Conspiracy to attack: malicious node collusion between slander nodes outside the gang, raise the evaluation of associates.

(4) Sleep attacks: Node If higher trust value is obtained within a certain period of time T has not to provide services, the trust value will always be maintained at a high level, which will cause the node to receive service only, without providing the service.

In Trust-BT behavior of nodes into trading behavior and feedback behavior. The transaction behavior refers to a node probability willing to provide the service, global trust value. Feedback behavior is the node evaluation to provide services. Slander and conspiracy attack node through feedback behavior, free-riding attack node achieved through trading behavior and sleep attacks.

Trading behavior, the nodes are divided into three categories: $\theta_i=1$ shows Altruistic Nodes (AN), at any time at 100% probability for other nodes to provide a good service. $\theta_i=0$ represents the selfish node (SN), at any time do not provide services or provide false malicious service. $0<\theta_i<1$ is mixed node (MN) selfishly profit big from time to time is selfish, the selfless acts profit is selfless.

Feedback behavior according to the node, the node is divided into two main categories: honest nodes and malicious nodes.

Honest node always accurately record transactions node behavior.

Malicious node is divided into simple malicious node NM (Naive and Malicious), the malicious collective CM (Collusive and Malicious), Strategy malicious node SM (Strategic and Malicious) node.

Simple malicious nodes are randomly distributed in the network, to defame honest trading behavior through feedback evaluation.

Malicious collective is more malicious nodes to form alliances through the feedback acts of nodes outside the gang slander, and improve feedback evaluation of the nodes within the gang.

Policy malicious nodes: These nodes first a selfless node identity to provide a good service, and other the elevated trust value slander selfless node

Journal of Theoretical and Applied Information Technology
15th November 2012. Vol. 45 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

elevation selfish nodes, and began to provide malicious services.

The trust model in the past that the low value of global trust node will slander and conspiracy attack, high trust value of the node is bound to provide honest feedback. It is noteworthy that a trusted node transaction behavior is not always truthful evaluation of other nodes, such as to provide a good service to 100% probability of selfless node does not necessarily honest nodes literature [6-7] to be confused with so vilified easier with the collusion attack.

This paper adopts the Bayesian discrimination method will filter out malicious node, honest node feedback evaluation estimated the true value of the trust. Bayesian discrimination is to accept the greater probability assumptions with Bayesian statistical inference theory [8-9] is consistent.

The probability to provide services so honest collection of nodes A evaluate the node i θ_0 collection B assessment egress i probability to provide services for θ , set the following two assumptions:

$$H_0 : \theta \in [\theta_0 - \varepsilon, \theta_0 + \varepsilon]$$

$$H_1 : \theta \notin [\theta_0 - \varepsilon, \theta_0 + \varepsilon]$$

ε is selected as a small positive number, such that $[\theta_0 - \varepsilon, \theta_0 + \varepsilon]$ and $\theta = \theta_0$ are difficult to discern.

Let θ_i be the probability of successfully providing service node i honest collection of nodes A and node i n transactions samples were collected in the time period T, where x equals the number of successful transactions.

Probability function of the number of successes X:

$$\pi(\theta_i | x) = \binom{n}{x} \theta_i^x (1-\theta_i)^{n-x}$$

Take θ_i priori distribution yoke priori $Beta(a,b)$ posterior probability density:

$$p(\theta_i | x) = \frac{\Gamma(n+a+b)}{\Gamma(x+a+b)} \theta_i^{x+a-1} (1-\theta_i)^{n-x+b-1}$$

Obtained the Bayesian estimation of θ_i is $\theta_0 = E[\theta_i] = \frac{a+x}{a+b+n}$.

Trading samples collected in time T cycle B provides a collection of k nodes of node i, wherein z represents the number of successful transactions, the probability function of the number of successful z:

$$f(z|\theta_i) = \binom{k}{z} \theta_i^z (1-\theta_i)^{k-z}$$

θ_i prior distribution for total conjugated priori $Beta(a,b)$ posterior probability density:

$$g(\theta_i | z) = \frac{\Gamma(k+a+b)}{\Gamma(z+a+b)} \theta_i^{z+a-1} (1-\theta_i)^{k-z+b-1}$$

Let $a_0 = \int_{\theta_0-\delta}^{\theta_0+\delta} g(\theta_i|z) d\theta_i$,

$a_1 = \int_0^{\theta_0-\delta} g(\theta_i|z) d\theta_i + \int_{\theta_0+\delta}^{1_0} g(\theta_i|z) d\theta_i$

If $\delta_1 = a_1 / a_0 < 1$ receiver H_0, that the feedback provided by the set B Evaluation and the similarity of the set A large, can be classified as honest nodes collection; if $a_1 / a_0 > 1$, then reject H_0, that in the honest A collection case, the set B of the evaluation of the probability is very small, that set B provide feedback evaluation exists the slander or collusion; if $a_1 / a_0 \approx 1$, continue to collect samples, to enter the next round of the hypothesis testing.

Probability and Statistics with respect to the classical method, Bayes hypothesis testing methods are simple and feasible, and do not need to select the hypothesis test statistic to determine the sampling distribution do not need to implement given significance level to determine the rejection region, and is easily extended to multiple hypothesis testing case, the classical hypothesis testing for more than three hypothetical question is difficult to deal with.

Sleep attacks, taking trust values the aging strategy. n_i represents the number of node i in the time period $T-1$ to provide a successful service, the lower limit of n_{low} set for each node within each cycle the number of successful transactions, if $n_i < n_{low}$, then the trust value attenuation in accordance with the next equation:

$$\theta_j(T) = \theta_j(T)(1 - e^{-\frac{n_i}{n_{low}}})$$

Inferred from the above formula, the fewer the number of transactions in the current period, aging faster decline in trust value of the faster. The more

Journal of Theoretical and Applied Information Technology
15th November 2012. Vol. 45 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the number of successful trading, trust the slower decline in value.

Free-riding behavior Trust-BT accurate inferred trust value of a node, each node with high trust value transactions, so that the selfish nodes have been checked.

4. SIMULATION

For the analysis of the proposed trust model, this paper established PeerSim [12] on the basis of the simulation environment. Simulation based on the Java programming environment. The program is run on a Pentium Dual-Core E5300 2.6GHZ, CPU, 2GB Memory standalone. The simulation application context file sharing, each simulation composed by a number of cycles T, each cycle, the node to find the desired file by a certain strategy and high node trust value is selected from all the node that owns the file download the file, empathy and the node high trust value of node upload files. Each node in each cycle for 50 queries, each query causes a transaction. The data collected at the end of each cycle, calculated into the next cycle. Each experiment was done 10 times and averaged. First cycle begins when the trust value of each node, the P2P scale, whichever is 1500 nodes, which selfless nodes with selfish nodes proportion 35% each, the mixed node accounted for 30%. Simple malicious node attacks, malicious collective attacks Policy malicious node attack case, the trust model Eigentrust and PStrust classic was conducted comparing the actual network environment are much more complex, more complex type of malicious nodes. Each experiment was done 10 times take the average, were divided into 4 groups.

Experimental evaluation criteria: download success rate the SDR (Successful Download Rate), the SDR = number of successful trading / total number of transactions. The when Trust-BT do Eigentrust to take selfless node SDR contrast, intuitively reflect the effect of the trust model because of the selfless node SDR changes.

Experiment 1: Free-riding attack

In the case of the introduction of 30% simple malicious nodes, the results of tests carried attack on free-riding is shown in Figure 1: selfless node AN SDR has been maintained at a high level, the selfish node SN SDR sharply with time mixed node MN SDR beginning when declined, terminated in the lowering of the SDR the selfish behavior SDR gradually increased. Figure 1 illustrates the existence of two types of nodes in the MN and SN aggressive behavior of free-riding has been checked.

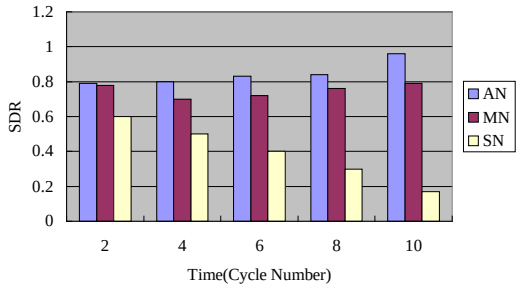

Figure 1: The Three Types Of Nodes SDR Changes Over Time

Experiment 2: NM class simulation

In the case of Figure 2 and Figure 3 is a simple introduction of different proportions of malicious nodes Nm node, the selfless node Trust_BT with Eigentrust SDR in comparative analysis of the results Fig. Randomly distributed in simple malicious node when both quite effective, at 50%, is still close to 75 percent of the SDR, as shown in Figure 2. But Eigentrust credibility larger the higher the trust value of node feedback rating. E.g. trust 0.8 nodes for the node provides feedback evaluation was 0.8, and i is the true value of 0.2, the trust value of the feedback provided by the node k of 0.2 was evaluated as 0.2, but Eigentrust that node j of the evaluation is more credible, resulting the estimated value of i is more deviated from the true value. If NM all nodes served by the high node trust value can be seen from Figure 3, the Trust-BT shows a clear advantage.

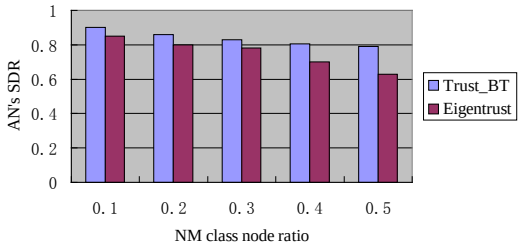

Figure 2: AN SDR In NM Change As Different Scale

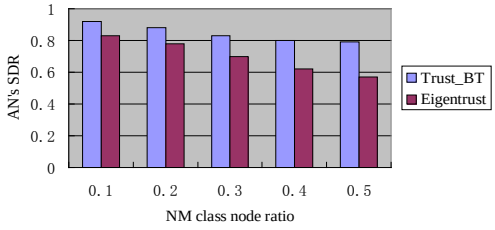

Figure 3: High Trust Value Node Acts As NM AN The SDR Changes With Different Proportion NM

Journal of Theoretical and Applied Information Technology
15th November 2012. Vol. 45 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Experiment 3: CM class simulation

In the case of Figure 4 is to introduce a different proportion of malicious collective CM node, the selfless node Trust-BT with Eigentrust SDR in comparative analysis of the results Fig. When collusion node uniform distribution of the various types of nodes in the network can be seen from Figure 4, for different sizes of CM Trust-BT model effectively suppress the collusion attack, the 50% CM circumstances, selfless node SDR still reached about 71%.

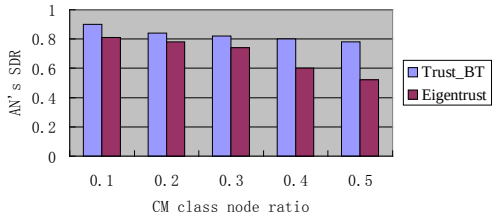

Figure 4: AN's SDR With CM Proportional Change

Experiment 4: SM attack

Policy malicious nodes is fixed at 30%, compare such case both systems selfless node SDR changes. Be seen from Figure 5, Eigentrust selfless node of the SDR decreases first slowly increased. While the selfless node Trust-BT SDR has maintained at a relatively high level. This is because the each SM node in the continuous selfless service for the system to get to the higher value of the trust, they slander selfless node and start the malicious transaction behavior. Eigentrust in nodes with high trust value the greater weight feedback evaluation, so selfless node trust value decline soon, SDR soon dropped, but due to the malicious transaction behavior of the SM node to make its own trust value decreased, feedback evaluation The weight to be diminished. Selfless node slowly restored high SDR. While in Trust-BT trust value calculated by the Bayesian discrimination method excludes malicious nodes feedback evaluation, selfless node trust value is always maintained at a higher level.

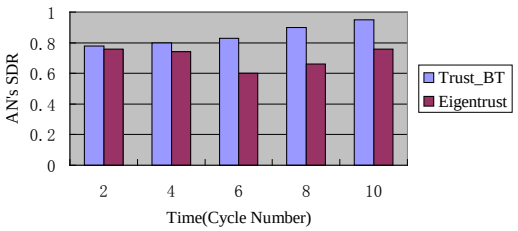

Figure 5: AN's SDR Changes Over Time

5. CONCLUSIONS

This paper studies the computational problems of trust between nodes in the P2P network analysis based on the projected global trust value evaluation model recommended by the local trust iteration, a new based on Bayesian theory trust model Trust-BT full use of a priori information and feedback evaluation sample, figure out a more accurate value of global trust, experimental simulation under various attacks. The mathematical analysis and simulation results show that, Trust-BT effectively isolate the malicious nodes and selfish nodes to improve the trading success rate of the entire network. Worthy of further study of this model is to improve slightly, they can apply to the wireless sensor network trust model.

REFRENCES:

[1] Michal Feldman, John Chuang, "Overcoming free-riding behavior in Peer-to-Peer Systems", ACM SIGecom Exchanges, Vol.5, No.4, 2005, pp. 41-50.

[2] Kudtarkar A M, Umamaheswari S, "Avoiding white washing in P2P networks", *Proceedings of the 1st International conference on Communication systems and networks*, IEEE Conference Publishing Services, March 23-25, 2005, pp. 1-4.

[3] Ganesh Kumar M, Arun Ram K, Ananya A R, "Controling free riders in Peer to Peer networks by intelligent mining", *Proceedings of the International conference on computer engineering and technology*, IEEE Conference Publishing Services, January 23-25, 2009, pp. 267-271.

[4] Kamvar S D, Schlosser M T, "EigenRep: Reputation Management in P2P Networks", *Proceedings of the 12th International world wide web conference*, IEEE Conference Publishing Services, May 20-24, 2003, pp. 123-134.

[5] Li Xiong and Ling Liu, "PeerTrust: Supporting Reputation-Based Trust for Peer-to-Peer Electronic Communities", IEEE transactions on knowledge and data engineering, Vol.16, No.7, 2004, 843-857.

[6] Song SS, Hwang K, Zhou RF, Kwok YK, "Trusted P2P transactions with fuzzy reputation aggregation", IEEE Internet Computing, Vol.9, No.6, 2005, 24-34.

[7] Wang YF, Y and Sakurai K, "Characterizing and reputation economic and social properties

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

of trust and reputation systems in P2P environment", Journal of computer science and technology, Vol.23, No.1, 2008, 129-140.

[8] Cox D R, Hinkley D V, Theoretical Statistics, Chapman-Hall, London, British, 1974.

[9] Saurabh Ganeriwal, Laura K Balzano, and Mani B Srivastava, "Reputation-based framework for high integrity sensor networks", ACM transactions on sensor networks, Vol.4, No.3, 2008, 1-15.

[10] Crosby G V, Pissinou N, "Cluster-based reputation and trust for wireless sensor networks", *Proceedings of the 4th Consumer Communications and Networking conference*, IEEE Conference Publishing Services, January 11-14, 2007, pp. 604-608.

[11] Audun Josang, "The Beta reputation System", *The 25th Bled electronic Commerce Conference e-Reality: Constructing the e-economy*, IEEE Conference Publishing Services, June 17-19, 2012, pp. 1-14.

[12] The PEERSIM website. [Online]. Available: http://peersim.sourceforge.net/2009-4-22.