Journal of Theoretical and Applied Information Technology
30th November 2012. Vol. 45 No.2

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

DYNAMIC TOPOLOGY ALGORITHM FOR P2P NETWORKS

[1]YUNXIA PEI, [2]JUANJUAN LUO

[1]Department of Math and Computer Science, Zhengzhou University of light Industry, Zhengzhou 450002, Henan, China

[2]Department of Economic Trade Management, Hunan Railway Professional Technology College, Zhuzhou 412001, Hunan, China

ABSTRACT

In order to improve the Peer-to-Peer network resource allocation efficiency, and to isolate the malicious nodes, this paper has presented a new method named dynamic topology algorithm (DTA), suitable for unstructured P2P network. As the neighbor node sets variable time-to-live (TTL) values, and in the interaction according to the neighbor node forwarding on the TTL value is adjusted; records of the query path directly on a forwarding node connection; given the forwarding information placing method and topological adjustment algorithm. The simulation experiments and analysis show that, the DTA topology construction method can be malicious nodes isolated to the network edge, improves the efficiency of resource locating, effectively improve the quality of service in P2P network.

Keywords: *Peer-to-Peer, Unstructured, Topology, Resource Locating*

1 INTRODUCTION

In recent years, the application of Peer-to-Peer (P2P) network increases rapidly, the P2P system allows a large number of cyber sources to be shared, such as file, calculation ability and so on. The P2P system is essentially a kind of distributed system, no central server. The P2P system in the node has interconnected to form a self-organizing overlapping network [1]. P2P is located in the underlying physical network topology and form a virtual overlay layer, for all types of P2P application support. The P2P network topology and P2P system operation are closely related, the influence of P2P network service quality. Some research shows that the P2P network topology optimization can improve the resource positioning efficiency, reduce the expense of the system, so as to improve the service quality in P2P network.

According to the topological structure of different P2P system is mainly divided into two categories: unstructured P2P systems and structured P2P system. At present, most of the popular P2P application is based on unstructured P2P systems. Unstructured P2P systems to random topological structure based on flooding and query, resulting in a large number of redundant network query message, affects the efficiency of resource locating [2]. In unstructured P2P network, a "good" topology can be more effective for locating resources, including the possibility of finding target service node, replication of the query message number, isolate the malicious nodes etc. The literature [3] at all angles of topology optimization technology, but did not consider network malicious node problem. The literature [4] node reliability adjusted according to topology, but to the neighbor node capacity calculation method is one-sided, loss of topology adjustment fairness. In order to improve the efficiency of resource locating, and isolate the malicious nodes, this paper presents a suitable for unstructured P2P network dynamic topology construction method DTA, to the neighbor node sets variable TTL value, and in the interaction of TTL value is adjusted, also recorded the query path directly on a forwarding node connections, and are based on these information to adjust the topology. The simulation experiments and analysis show that the topology construction method DTA is effective. It can be malicious node isolation to the edge of the network, improve the service quality.

2 THE RELATED WORK

Fast Track[5] is based on the architecture of P2P protocol. The protocol has selected high performance of node as the super node to other nodes, as a common node. The introduction of the portion relative ability of the node as the super node bearing more computing work and in order to accelerate the speed of query routing. The literature

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

[6] is based on node in all aspects of the processing ability of topology optimization method, respectively, according to the node degree, routing or multiple measuring standard comprehensive to select neighbor node. The existing P2P network topology structure and the underlying physical network topology often present the topology mismatching problem, the literature [7] for the topology mismatch problem, put forward a kind of position related topology matching technology and structural topology algorithm, which selects the physical location is closer to the node as a neighbor, but this method requires all node clock is synchronized with the. Adaptive Connection Establishment (ACE) for each node to its neighbors by establishing the minimum spanning tree, optimization of neighbor nodes, can dynamically connect to a physical location closer to the node. The method solves the topology mismatch problem, and the decrease of P2P network traffic. The P2P network nodes in the content of interest often show a certain concentration, so some research into peer interest structure topology. The literature [8] is how to use the user's interest to form different clusters. This paper presents a dynamic adaptive algorithm and look-up method. The literature [9] in the existing network topology of the above, proposed one kind based on the node of interest shortcut algorithm, nodes in initiating the query, first sent to the shortcut search, until if in shortcuts is not responding to, will continue in the P2P network topology sends a query, but from the shortcut of response is not necessarily the most reliable response. The literature [10] provides a P2P network with a similar interest in the nodes form a "cluster" method, can reduce the search range query request, in a short distance to get response.

The topology optimization method does not consider the problem of malicious nodes. Based on a kind of node reliability topology optimization method based on the node to provide the service of confidence degree adjustment topology, nodes are more likely to choose high trust value node as neighbors, to provide reliable service quality of malicious nodes have smaller chance of being selected as the neighbor. The literature [11] proposed credibility sensitive topology evolution algorithm, based on the node trust value determines the topology hierarchy. APT for a class of hidden malicious nodes proposed topology construction protocol, but the query only record the first forward the query to the neighbor node, and according to the final results returned for the first neighbor node is disposed, is unfair, will lead to the normal node. RC-ATP is also the record corresponding to the response of the first forwarding of neighbor nodes, different nodes from their own interactive experience to calculate the other node through mutual reciprocal capacity, and ability to measure other nodes, and on this basis to select neighbor nodes, but about the neighbor node forwarding capacity calculation method of one-sided.

Literature [12] based on profit sharing model topology construction, adjustment of each node in the measure of using the service reliability and credibility are weighted and routing, node in the calculation of each subsequent neighbor node routing reliability, integrated the node follow-up neighbor node from the node through the forward all the query get a share of the profits, computational overhead is large, in addition, in determining the source node based on the feedback credibility when public interactive node algorithm, in a large-scale network trading sparse cases are difficult to determine, thus giving routing reliability judgment brought difficulty. Put forward a kind of Gnutella system P2P network file sharing protocol Adaptive Gnutella-like Protocol (AGP), through the node for service contribution assessment of node reliability, according to the credibility and share content similarity of nodes in the network will form a "cluster", collaboration of good node and credibility of the nodes become topology core.

3 DTA RESOURCE LOCATOR

P2P networks can be expressed as an undirected graph name $G = (P, E)$, where P is the set of nodes, E is the edge set, (i, j) said node i and node connection to j, $i, j \in P$. P2P network connection is symmetrical, can be used to describe a neighbor node set, the neighbor node is bidirectional, and node i neighbor nodes can be expressed as $N(i) = \{ j \mid (i, j) \in E \}$.

Firstly it is introduced the DTA basic resource location method. Nodes can be similar to Gnutella network mechanisms for processing nodes join and query, i.e. node with some currently at the nodes of the network connection is established to get connected to the Internet, and through the flooding way to deliver in the network query message, and through TTL to control query size, TTL to express the message in the network the transmission frequency, each node in the query message to the neighbor node transfer before the TTL minus 1, when TTL is zero, no longer passes the query message. In P2P file sharing network as an example, the resource location method is described as follows:

Journal of Theoretical and Applied Information Technology
30th November 2012. Vol. 45 No.2

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

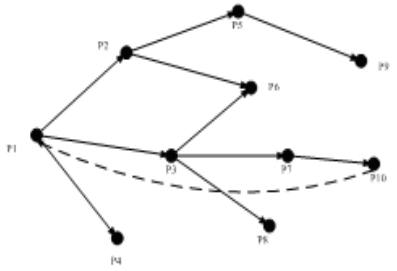

Figure 1: Resource Location Sketch Map

In Figure 1, the node P1 as the requesting node, a query message to the neighbor node, neighbor node P2, P3 and P4 respectively the query message to their respective neighbor nodes, until TTL 0. Node P1 received a series of response, and from the response list in a mechanism to choose the response node P10 and sends request message, wait for the node P10 to provide services, in response to the selection mechanism determines the node priority with which node to trade. The node P1 resource locating process, query path for P1-P3-P7-P10, P3 and P7 for forwarding node, P3 for the requesting node P1 neighbor node, P7 known as correspond to the service node P10 direct forwarding node. Query path generally have multiple forwarding node, but compared to the other node, direct forwarding node in the forwarding path plays a crucial role in DTA, therefore, in addition to record query path on the neighbor node forwarding results, also recorded the forwarding node connection, and based on these information to adjust the topology.

4 FORWARDING INFORMATION POSITION

In the DTA topology construction method in P2P network, each node corresponds to a neighbor node list and a connecting factor list, as shown in Table 1 and Table 2 shows. Neighbor node list recorded in each neighbor nodes and the corresponding ID trusted TTL value; connecting factor list records corresponding to the service node forwarding node connection factor, representing a direct forwarding node connection service node capacity, reflecting the direct forwarding node connection situation. Node i neighbor list as shown in Table 1, n_k said the neighbor node corresponding to the ID; $TTL^i_{n_k}$ for node i, the neighbor node n_k corresponding trusted TTL value. Node i connection factor list as shown in Table 2, said t_k ID $f^i_{t_k}$ said forwarding node; direct forwarding node t_k connection factor; $Succ^i_{t_k}$ expressed through the node t_k directly forward, the node i success response times; $Fail^i_{t_k}$ expressed through the node t_k forwarded directly, node i get the failure response times. Wherein the connecting factor is defined as follows:

Definition 1: For the node i, node t_k connection factor

$$f^i_{t_k} = \frac{f^i_{t_k} \times \left(Succ^i_{t_k} + Fail^i_{t_k}\right) + TranSat^i_{t_k}}{Succ^i_{t_k} + Fail^i_{t_k} + 1} \quad (1)$$

In the formula (1), $f^i_{t_k}$ said node i on direct forwarding node t_k previous connection in the comprehensive evaluation; $TranSat^i_{t_k}$ forwarding node t_k for the current interactive contribution, the current interaction is successful, for interactive satisfaction of 1/2 or -0.5.

Table 1: Neighbor List

Neighbor ID	Trusted TTL
n_1	$TTL^i_{n_1}$
n_2	$TTL^i_{n_2}$
…	…
n_k	$TTL^i_{n_k}$

Table 2: Connection Factor List

Transmit Peer ID	Connection factor	Success Time	Failure Time
t_1	$f^i_{t_1}$	$Succ^i_{t_1}$	$Fail^i_{t_1}$
t_2	$f^i_{t_2}$	$Succ^i_{t_2}$	$Fail^i_{t_2}$
…	…	…	…
t_k	$f^i_{t_k}$	$Succ^i_{t_k}$	$Fail^i_{t_k}$

5 TOPOLOGY ADJUSTMENT ALGORITHM

With a group of first topology construction method DTA will use several parameter definition.

$Nb(i)$: Node i neighbor node set;

$Nb_{\max}$: Node can be connected up to the neighbor number;

$Min(S)$: Set S accumulated trust the lowest node;

$Max(S)$: Set S in the node with the highest cumulative trust;

Journal of Theoretical and Applied Information Technology
30th November 2012. Vol. 45 No.2

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

$FMax(S)$: Set S in the node with the highest cumulative connection factor;

$Num(S)$: Set S the number of nodes in;

$TTL(i,n)$: Node i neighbor node, a trusted TTL value for the n neighbor node set;

$Local(i)$: For the node i to provide the service nodes collection;

$Tran(i)$: For the node i to provide query forwarding and go directly to the response set of nodes, node i record query path corresponding to the service node directly forwarding node, namely in the query to the service node until after the last node;

j_d : The serving node j direct forwarding node;

$SumTrust(i, j)$: i appears in node, node j good cumulative trust value;

$SumTran(i, j)$: i appears in node, node j good cumulative connection factor $f_j^i * (Succ_j^i - Fail_j^i)$.

Related primitives are as follows:

Disconnect(i, j): Node i and node j connection;

AddNeighbor(i, j): Node i and node j to increase mutual neighbor;

RequestConnect(i, j): Node i request and j as the neighbor node, if the request is accepted, the return of *true* , *false* otherwise;

ConnectToRandomPeers(i): Node i in the network with random connections a neighbor node;

The node at each interaction has completed, need to update the local service node information, information of nodes and connecting factor information recommendation. When reaching the topology adjustment time, according to the degree of satisfaction with the service node, the implementation of different adjustment procedure.

5.1 Unsatisfied With Service Of Topological Adjustment

If node i for service node j not satisfied with the services provided, need to adjust the neighbor node and its TTL value, adjust the algorithm is as follows:

1) The requesting node i update query path directly on a forwarding node connection factor;

2) According to the service request node i corresponding to the node j hops $Hops_j^i$ ($Hops_j^i$ said node i sends the query message arrives at the service node j has the transmission frequency) to update the query path corresponds to the neighbor node n_k is the credibility of $TTL_{n_k}^i = Hops_j^i - 1$;

3) If $TTL_{n_k}^i$ is 0, or a service node j for the requesting node i neighbor node, node i broken and node j connection, and then connected with a new neighbor node.

4) If $TTL_{n_k}^i$ is 1, indicating a request to the neighbor node i at n_k trusted TTL is 1, then the neighbor node will no longer have the forwarding function, only as a response to the node, so the request node need to determine left the neighbor node is worth it. If the neighbor nodes' trust values and cumulative trust value satisfies the condition, and the neighbor node TTL is equal to 1 of the neighbor node number insufficient neighbor node total 1/2, can leave the neighbor node; if not less than half of the total number of neighbor nodes, are available in the neighboring nodes to select a replacement for comparison, the neighbor node's cumulative trust value and the rest of the TTL value of 1 neighbor node in the lowest cumulative trust value, if higher than the existing neighbor minimum cumulative trust value is broken, the cumulative trust value lowest neighbor node; otherwise break the neighbor node, and then reconnect a new neighbor node. Specified in DTA TTL 1 neighbor node number no more than half of the total number of neighbor nodes, is to ensure that the query proliferation and response number.

5) Other, the updated TTL value at least two level forwarding, operation is not conducted, return;

Unstructured P2P system based on flooding query diffusion through TTL to control scale, according to the neighbor node forwarding, for each neighbor node assigns different trusted TTL value, not only to ensure that the query diffusion, but also reduce the network query message flow. The connection of a new neighbor node method is as follows: in local records history information service node selects trusted service node, and the connecting factor no negative forwarding record, according to their cumulative trust value from high to low one by one in the order request. If the connection requests are not accepted, then look at the connecting factor list and not in the local service node information in the node, select reliable forwarding node, according to the accumulated connecting factor from high to low to send the connection request. If still no find is a randomly chosen node is connected to the connection request is accepted.

Journal of Theoretical and Applied Information Technology
30th November 2012. Vol. 45 No.2

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

5.2 Service Satisfaction Of Topological Adjustment

The requesting node i first update query path directly on a forwarding node connecting factor, then check the connection factor list service node connecting factor is whether the negative forwarding record, if not negative forwarding record, is trying to the service node for the neighbor.

1) If not his neighbor, and when the number of neighbor nodes is less than $Nb_{\max}$: first try to put the service node j as neighbors, if the connection request is received, the node j as neighbors; or try to connect again the forwarding node j_d , the node j_d will need to meet two conditions: no negative service record; no negative transmit records.

2) If it is not a neighbor, and when the neighbor node number is not less than $Nb_{\max}$, then from the neighbor list to find a cumulative the trust value lowest neighbor node $Min(Nb(i))$ to prepare replacement. If the neighbor node $Min(Nb(i))$ cumulative trust value less than the satisfaction threshold or less than the cumulative trust value of the service node j, try to request the connection service node j, if the connection request is accepted, then cut off a neighbor node $Min(Nb(i))$; if the connection request is not accepted, try to connect directly corresponds to the response forwarding node j_d , the node j_d also need to meet the two conditions: no negative record of service; no negative forwarding record.

5.3 Receiving Connection Request

Receive connection requests node j requires judgment, and then decide whether to accept the connection request.

1) Node j local service node information recording sheet and connecting factor list, if there are negative service record or negative node i initiate connection requests forwarding record, do not accept the connection request of node i;

2) If there is no negative record node j View the number of neighbors, neighbors number not reached the maximum $Nb_{\max}$, then accept the connection request; neighbors reached the maximum number $Nb_{\max}$, the existing neighbors looking for the lowest trust value node in the replacement, if the value is less than the trust value of node i trust the neighbors, to accept the connection request.

6 SIMULATION AND RESULTS ANALYSIS

6.1 Simulation Environment

In this paper, a query cycle model simulation, and construct a P2P file-sharing networks. Each simulation consists of a number of simulation cycles, in each simulation cycle, a node in the network can initiate the file query and respond to queries, the query request broadcast manner similar Gnutella's through the TTL control the size of the query. Node receives a query message to see if its response, initiated query node waits to receive a response from the response list to select the highest trust value node download until the transaction is successful or all responses were unsuccessful. In the simulation, the nodes in the network are divided into two categories: normal nodes and malicious nodes; normal node to provide reliable service, malicious nodes provide reliable service.

6.2 DTA Methods To Assess

In this paper, we have used the characteristic path length to assess DTA method. Node i average characteristic path length $cpl_i = \frac{1}{|P \setminus i|} \sum_{j \in P \setminus i} shortestPath(i, j)$, i for node and other nodes in the network the shortest path between the average value, P / i said the network contains node i, other nodes collection. When a node and other nodes between paths does not exist, it is the shortest path for the 15.

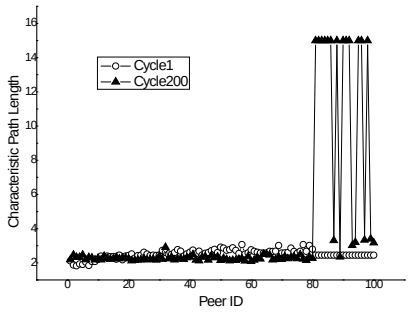

Figure 2: Each Node To The Nodes Of The Average Path Length

Figure 2 shows the first query the cycle before and 200th query cycle, each node to all other normal cooperative node average characteristic path length. Every node 1 to 80 is normal node, for the rest of the malicious nodes. As you can see from Figure 2, in the first query cycle topology adjustment before, node average characteristic path length difference, around 2 and 3; in 200th query cycle topology adjustment, can be seen after several interactive, normal node to the other node of the normal average

Journal of Theoretical and Applied Information Technology
30th November 2012. Vol. 45 No.2

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

characteristic path did not change significantly, and slightly lower, and most of the malicious nodes to normal node average characteristic path length is greatly increased.

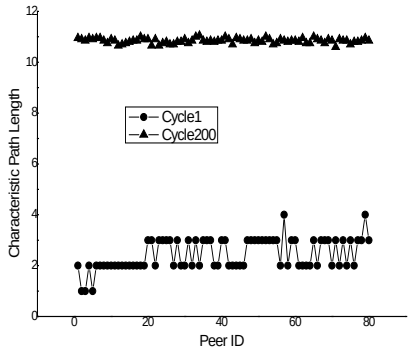

Figure 3: Normal Node To The Other Malicious Node Average Characteristic Path Length

Figure 3 shows the first query and a 200th query cycle period, normal cooperative node to all other malicious node average characteristic path length. Normal node with the query cycle increased, to other malicious node average characteristic path length increasing, reducing the malicious nodes to provide reliable service opportunities.

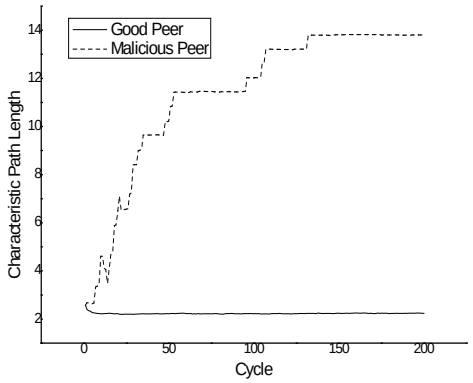

Figure 4: Node Average Characteristic Path Length With The Periodic Variation

Figure 4 shows the normal cooperative nodes and the malicious node to all other normal cooperative node average characteristic path length with the query cycle changes. Figure 4 shows that the normal node to other nodes, the average characteristic path length is very stable, maintained at a relatively low value; and the malicious nodes to normal node average characteristic path length with the interactive cycle increased, the malicious nodes gradually excluded to the edge of the network.

Figure 5 shows the normal cooperative node and malicious node average number of neighbors with query cycle changes. Figure 5 shows that the average number of neighbors, the normal node at the beginning of the polling period is increased continuously, and then maintained at can have the maximum number of connections and malicious nodes near; average number of neighbors with the polling period increase decrease was constant, this is due to malicious nodes to provide reliable service, more and more normal node broke and malicious node connection.

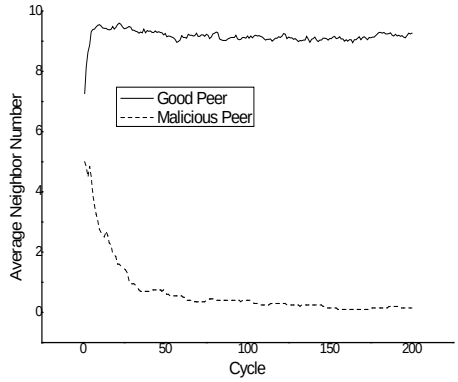

Figure 5: The Average Number Of Neighbors With Query Cycle Changes

With similar interests and the quality of service nodes of a cluster or neighbors, these clusters will be on the malicious node unreachable. In the DTA protocol, the connection between is based on previous download results, therefore, the more similar the interest between nodes will be more interactive connection is established, the greater the probability that the nodes can, through the shorter route of access to services, this phenomenon is also called as "small world phenomenon". Node i clustering coefficient is defined as $C_i = \frac{2E_i}{k_i(k_i-1)}$, the number of neighbors for node is k_i, E_i for these neighbor nodes are connected between side. This paper assesses the in each cycle the average cluster coefficient. Figure 6 shows that the average clustering coefficient, the node with the polling period increase which increases constantly, interest in similar nodes form a cluster, interact with each other more closely.

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

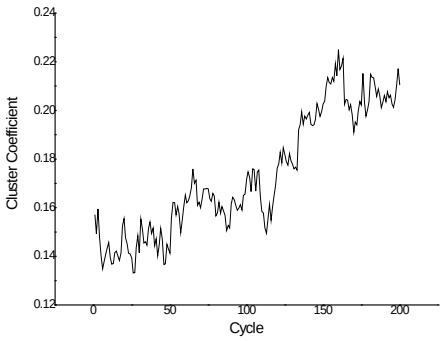

Figure 6: Average Clustering Coefficient With Periodic Variation

7 CONCLUSIONS

Topology adjustment DTA considered the connected factors and the neighbor node's TTL value, and APT topology algorithms, despite increasing the number of storage overhead, but topological adjustment cost reduction. The simulation experiments and analysis show that the topology construction method DTA validity, topology to the smaller overhead positioning reliable resources, and the malicious node isolation to the edges of the network, improve the service quality.

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 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.