Report on Reputation Based Data Aggregation For Wireless Network

Bhoopendra Singh M Tech (CSE) [1] , Mr. Gaurav Dubey[2]
[1]Amity University ,Noida, India
[2]Amity School of Engineering and Technology, Noida, India

ABSTRACT:
wireless sensor network consists of spatially distributed autonomous sensors to monitor physical or environmental conditions, such as temperature, sound, pressure, etc. and to cooperatively pass their data through the network to a main location.In wireless networks, malicious sensor nodes send false data reports to distort aggregation results. Existing trust systems rely on general reputation to mitigate the effect of this attack. This report is the implementation of one of a novel reliable data aggregation protocol, called RDAT i.e Reliable Data Aggregation Protocol.In this report Reliable Data Aggregation Protocol with functional reputation is implemented. It is based on the concept of functional reputation. Functional reputation enables data aggregators to evaluate each type of sensor node action using a respective reputation value thereby increasing the accuracy of the trust system. The simulation results show that protocol RDAT significantly improves the reliability of aggregated data in the presence of compromised nodes.

I. INTRODUCTION

The field of wireless sensor networks combines sensing, computation, and communication into a single tiny device. Through advanced mesh networking protocols, these devices form a sea of connectivity that extends the reach of cyber space out into the physical world. As water flows to fill every room of a submerged ship, the mesh networking connectivity will seek out and exploit any possible communication path by hopping data from node to node in search of its destination. The power of wireless sensor networks lies in the ability to deploy large numbers of tiny nodes that assemble and configure themselves. Usage scenarios for these devices range from real time tracking, to monitoring of environmental conditions, to ubiquitous computer environment, to in situ monitoring of the health of structures or equipment. The application demands for robust, scalable, low-cost and easy to deploy networks are perfectly met by a wireless sensor network. If one of the nodes should fail, a new topology would be selected and the overall network would continue to deliver data. If more nodes are placed in the field, they only create more potential routing opportunities. There is extensive research in the development of new algorithms for data aggregation, ad-hoc routing and distributed signal processing in context of wireless sensor networks. As the algorithms and protocols for wireless sensor network are developed, they must be supported by a low power, efficient and flexible hardware platform.

1.1 Overview of wireless sensor network

The concept of wireless sensor networks is based on a simple equation:

Sensing+CPU+Tranceiver=Thousands of potential applications

As soon as people understand the capabilities of a wireless sensor network, hundreds of applications spring to mind. It seems like a straightforward combination of modern technology. However , actually combining sensors , radios and CPUs into an effective wirelss sensor network requires a detailed understanding of the both capabilities and limitations of each of the underlying hardware components, as well as detailed understanding of modern networking technologies and distributed systems theory. Each individual node must be designed to provide the set of primitives necessary to synthesize the interconnected web that will emerge as they are deployed, while meeting strict requirements of size, cost and power consumption. Recent advances in micro-electro-mechanical systems (MEMS) technology, wireless communications, and digital electronics have enabled the development of low-cost, low-power, multifunctional sensor nodes that are small in size and communicate in

short distances. These tiny sensor nodes, which consist of sensing, data processing, and communicating components, leverage the idea of sensor networks based on collaborative effort of a large number of nodes.
Sensor networks represent a significant improvement over traditional sensors, which are deployed in the following two ways :
• Sensors can be positioned far from the actual phenomenon, i.e., something known by sense perception. In this approach, large sensors that use some complex techniques to distinguish the targets from environmental noise are required.
• Several sensors that perform only sensing can be deployed. The positions of the sensors and communications topology are carefully engineered.

They transmit time series of the sensed phenomenon to the central nodes where computations are performed and data are fused. A sensor network is composed of a large number of sensor nodes, which are densely deployed either inside the phenomenon or very close to it. The position of sensor nodes need not be engineered or pre-determined. This allows random deployment in inaccessible terrains or disaster relief operations. On the other hand, this also means that sensor network protocols and algorithms must possess self-organizing capabilities. Another unique feature of sensor networks is the cooperative effort of sensor nodes. Sensor nodes are fitted with an on-board processor. Instead of sending the raw data to the nodes responsible for the fusion, sensor nodes use their processing abilities to locally carry out simple computations and transmit only the required and partially processed data.

1.2 Classification System Design

Classification plays a vital role in many information management and retrieval tasks. Based on the organization of categories, The sensor nodes are usually scattered in a sensor field as shown in Fig. 1. Each of these scattered sensor nodes has the capabilities to collect data and route data back to the sink and the end users. Data are routed back to the end user by a multi-hop infrastructure less architecture through the sink.The sink may communicate with the task manager node via Internet or Satellite. This protocol stack combines power and routing awareness, integrates data with networking protocols, communicates power efficiently through the wireless medium, and promotes cooperative efforts of sensor nodes. The protocol stack consists of the application layer, transport layer, network layer, data link layer, physical layer, power management plane, mobility management plane, and task management plane. Depending on the sensing tasks, different types of application software can be built and used on the application layer. The transport layer helps to maintain the flow of data if the sensor networks application requires it. The network layer takes care of routing the data supplied by the transport layer. Since the environment is noisy and sensor nodes can be mobile, the MAC protocol must be power aware and able to minimize collision with neighbors' broadcast. The physical layer addresses the needs of a simple but robust modulation, transmission and receiving techniques. In addition, the power, mobility, and task management planes monitor the power, movement, and task distribution among the sensor nodes. These planes help the sensor nodes coordinate the sensing task and lower the overall power consumption.

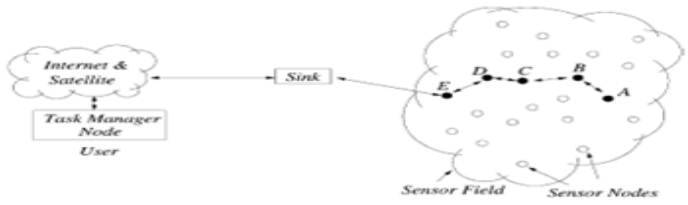

Fig. 1 : Sensor nodes scattered in a sensor field.

The power management plane manages how a sensor node uses its power. For example, the sensor node may turn off its receiver after receiving a message from one of its neighbors. This is to avoid getting duplicated messages. Also, when the power level of the sensor node is low, the sensor node broadcasts to its neighbors that it is low in power and cannot participate in routing messages. The remaining power is reserved for sensing. The mobility management plane detects and registers the movement of sensor nodes, so a route back to the user is always maintained, and the sensor nodes can keep track of who are their neighbor sensor nodes. By knowing who the neighbor sensor nodes are, the sensor nodes can balance their power and task usage. The task management plane balances and schedules the sensing tasks given to a specific region. Not all sensor nodes in that region are required to perform the sensing task at the same time. As a result, some sensor nodes perform the task more than the others depending on their power level. These management planes are needed, so that sensor nodes can work together in a power efficient way, route data in a mobile sensor network, and share resources between sensor nodes. Without them, each sensor node will just work individually. From the whole sensor

network standpoint, it is more efficient if sensor nodes can collaborate with each other, so the lifetime of the sensor networks can be prolonged.

II. APPROACHES

2.1 Functional Reputation Based Data Aggregation

Consider a large sensor network with densely deployed sensor nodes. Due to the dense deployment, sensor nodes have overlapping sensing ranges and events are detected by multiple sensor nodes. Hence, aggregation of correlated data at neighboring sensor nodes is needed. Some sensor nodes are dynamically designated as data aggregators to aggregate data from their neighboring sensor nodes, although every sensor node is assumed to be capable of doing data aggregation. To balance the energy consumption of sensor nodes, the role of data aggregator is rotated among sensor nodes based on their residual energy levels. Sensor nodes have limited computation and communication capabilities. For example, the Mica2 motes have a 4Mhz 8bit Atmel microprocessor, and are equipped with an instruction-memory of 128KB and a RAM of 4KB. All messages are time-stamped and nonces are used to prevent reply attacks. Sensor nodes employ monitoring mechanisms to detect malicious activities of their neighbours. Sensor nodes establish pairwise shared keys with their neighbours using an existing random key distribution protocols . Pairwise keys are used for data authentication. Data are transmitted in plain text unless it is stated otherwise. Intruders can compromise sensor nodes via physical capturing or through the radio communication channel. Once a sensor node is compromised, all information of the node becomes available to the intruder. Although compromised nodes can perform many types of attacks to degrade the network's security and performance, we only consider the attacks against integrity of the aggregated data. We assume that compromised nodes send false data (sensing reports) to data aggregators. If a compromised node is selected as data aggregator it can inject false data into aggregated data. In addition, compromised nodes selectively forward and misdirect aggregated data to distort the integrity of the aggregated data.

2.2 Reliable data aggregation protocol (RDAT)

The basic idea behind protocol RDAT is to evaluate trustworthiness of sensor nodes by using three types of functional reputation, namely sensing, routing, and aggregation .Sensor nodes monitor their neighborhood to obtain first-hand information regarding their neighboring nodes. For sensing, routing, and aggregation tasks, each sensor node Ni records good and bad actions of its neighbors in a table referred to as functional reputation table. Functional reputation tables are exchanged among sensor nodes to be used as second-hand information during trust evaluation. The functional reputation tables are piggy backed to other data and control packets in order to reduce the data transmission overhead. When sensor node Ni needs to interact with its neighbour Nk , Ni evaluates the trustworthiness of Nk using both first-hand and second-hand information regarding Nk . Functional reputation for aggregation ($R_{a,b}^{aggregation}$) is needed by sensor nodes to evaluate the trustworthiness of data aggregators. Functional reputations for routing ($R_{a,b}^{routing}$)and sensing ($R_{a,b}^{sensing}$) are used by data aggregators to increase the security and reliability of the aggregated data. Functional reputation values are quantified using beta distributions of node actions as explained next.

2.3 Beta reputation system

As the success of Bayesian formulation in detecting arbitrary misbehavior of sensor nodes is , we select a Bayesian formulation, namely beta reputation system, for trust evolution. In this section, before giving the details of protocol RDAT, we present a brief information about beta reputation system. Posteriori probabilities of binary events can be represented as beta distributions which is indexed by the two parameters α and β . The beta distribution f (p|α,β) can be expressed using the gamma function Γ as:

$$f(p|\alpha,\beta) = (\Gamma(\alpha+\beta)/\Gamma(\alpha)+\Gamma(\beta))\, p\alpha-1(1-p)\beta-1$$
$$0 \le p \le 1,\ \alpha > 0,\ \beta > 0$$

The probability expectation value of the beta distribution is given by $E(p) = \alpha/(\alpha+\beta)$. To show that how beta function can be employed in sensor networks let us consider the task of target detection as an action with two

possible outcomes, namely "correct" and "false". Let r be the observed number of "correct" target detections and s be the the observed number of "false" target detections by a sensor node. The beta function takes the integer number of past observations of "correct" and "false" target detections to predict the expected frequency of "correct" target detections by that sensor node in the future which is achieved by setting:

$$\alpha = r+1 \ \beta = s+1, \text{ where } r, s \geq 0.$$

The variable p represents the probability of "correct" target detections and f (p|α,β) represents the probability that p has a specific value. The probability expectation value is given by E(p) which is interpreted as the most likely value of p. Hence, a sensor node's reliability can be predicted by beta distribution function of its previous actions as long as the actions are represented in binary format.

2.4 Computing functional reputation and trust

Functional reputation value ($R_{a,b}^{X}$) is computed using beta density function of sensor node Nk's previous actions with respect to function X. Trust ($T_{i,j}^{X}$) is the expected value of $R_{a,b}^{X}$.Let us take routing task as an example. If sensor node Ni counts the number of good and bad routing actions of Nk as α and β, respectively. Then, Ni computes the functional reputation $R_{a,b}^{routing}$ about node Nk as Beta(α+1,β+1). Following the definition of trust, $T_{i,j}^{routing}$ is calculated as the expected value of $R_{a,b}^{routing}$

$$T_{i,j}^{routing} = E(Beta(\alpha+1,\beta+1))$$
$$= \alpha+1/\alpha+\beta+2$$

This equation shows that the expected value of the beta distribution is simply the fraction of events that have had outcome α. Hence, functional reputation value of routing is given by the ratio of good routing actions to total routing actions observed. This is an intuitive decision and it justifies the use of the beta distribution. In the above formula, $R_{a,b}^{routing}$ represents node Ni's observations about node Nk . In other words, it just involves first-hand information. Reputation systems that depend on only first-hand information has a very large convergence time . Hence, second-hand information is desirable in order to confirm firsthand information. In protocol RDAT, neighboring sensor nodes exchange their functional reputation tables to provide secondhand information and this information is included in trust evaluation. Let us assume that sensor node Ni receives secondhand information about node Nk from a set of N nodes and $S_{info}(r_{k,j})$ represents the second-hand information received from node Nk (k ∈ N). Ni already has previous observations about Nj as $\alpha_{i,k}$ and $\beta_{i,j}$. Further assume that, in a period of Δt, Ni records $r_{a,b}$ good routing actions and $s_{i,j}$ bad routing actions of Nk . Then, Ni computes the trust $T_{i,j}^{routing}$ for Nk as follows.

$$\alpha_{i,j}^{routing} = \nu*\alpha_{i,j} + r_{a,b} + \Sigma\, S_{info}^{routing}(r_{k,j})$$
$$\beta_{i,j}^{routing} = \nu*\beta_{i,j} + r_{i,j} + \Sigma\, S_{info}^{routing}(r_{k,j})$$
$$T_{i,j}^{routing} = E(beta(\alpha_{i,j}^{routing}+1, \beta_{i,j}^{routing}+1))$$

where ν < 1 is the aging factor that allows reputation to fade with time. Integration of first and second hand information into a single reputation value is studied in by mapping it to Dempster-Shafer belief theory . We follow a similar approach and use the reporting node Nk's reputation to weight down its contribution to the reputation of node Nk . Hence, second-hand information $S_{info}(r_{k,j})$is defined as

$$S_{info}(r_{k,j}) = (2*\alpha_{i,k} * r_{k,j})/((\beta_{i,k}+2) * (r_{k,j} + s_{k,j} + 2) * (2 * \alpha_{i,k}))$$
$$S_{info}(s_{k,j}) = (2*\alpha_{i,k} * s_{k,j})/((\beta_{i,k}+2) * (r_{k,j} + s_{k,j} + 2) * (2 * \alpha_{i,k}))$$

The idea here is to give greater weight to nodes with high trust and never give a weight above 1 so that second-hand information does not outweigh first-hand information. In this function, if $\alpha_{i,k} = 0$ the function returns 0, therefore node Nk's report does not affect the reputation update.

2.5 Secure and reliable data aggregation

In protocol RDAT, data aggregation is periodically performed in certain time intervals. In each data aggregation session, secure and reliable data aggregation is achieved in two phases. In the first phase, before transmitting data to data aggregators, each sensor node *Ni* computes $R_{a,b}^{aggregation}$ value for its data aggregator Aj and evaluate the trustworthiness of Aj . If trustworthiness of Aj is below a predetermined threshold, then Ni does not let Aj to aggregate its data. To achieve this, Ni encrypts its data using the pairwise key that is shared between the base station and Ni and sends this encrypted data to the base station along with a report indicating Ajmay be compromised. Based on the number of reports about Aj over the time, the base station may decide that Aj is a compromised node and it should be revoked from the network. In the second phase of data aggregation session, the following Reliable Data Aggregation (RDA) algorithm is run by data aggregators. Algorithm RDA depends on $R_{a,b}^{sensing}$ and $R_{a,b}^{routing}$ functional reputation values to mitigate the effect of compromised sensor nodes on aggregated data.

The Algorithm RDA is-

- Input: Data aggregator Aj , Aj's neighboring nodes {N1,N2, ...,Ni}, trust values of neighboring nodes computed by Aj { $T_{j,1}^{sensing}$,..., $T_{j,i}^{sensing}$ } and{ $T_{j,1}^{routing}$,..., $T_{j,i}^{routing}$ }.
- Output: Aggregated data Dagg .
- Step 1: Aj requests each Ni to send its data for data aggregation.
- Step 2: Sensor nodes {N1,N2, ...,Ni} transmit data {D1,D2, ...,Di} to Aj .
- Step 3: Aj updates trust values $T_{i,j}^{sensing}$ and $T_{i,j}^{routing}$ of each Ni based on the first and second hand information regarding Ni .
- Step 4: Aj weights data Di of sensor node Ni using the $T_{i,j}^{sensing}$ and $T_{i,j}^{routing}$.
- Step 5: Aj aggregates the weighted data to obtain Dagg.

Since compromised nodes send false sensing reports in order to deceive the base station, Algorithm RDA considers trustworthiness of sensor nodes with respect to sensing function to increase the reliability of aggregated data. To achieve this, Aj weights data of each sensor node Ni with respect to the sensor node's trust value $T_{i,j}^{sensing}$ and $T_{i,j}^{routing}$. By weighting sensor data based on trust levels, data aggregators reduce the compromised sensor nodes' effect on the aggregated data. This reason is that a compromised node Ni is expected to have low $T_{i,j}^{sensing}$ and $T_{i,j}^{routing}$ values as shown in next section.

3.1 Experimental Simulation and Results

These various algorithms have their implemented results upon which simulations have carried out in order tomeasure the performance parameters of the algorithms over the datasets. The results are summarized in the

following tables.

Fig – cluster formation

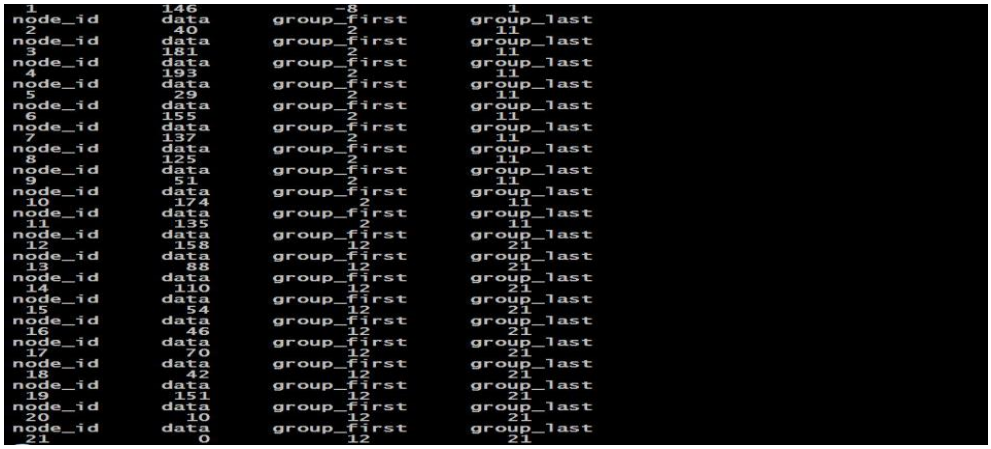

Fig- node data

Fig – functional reputation table

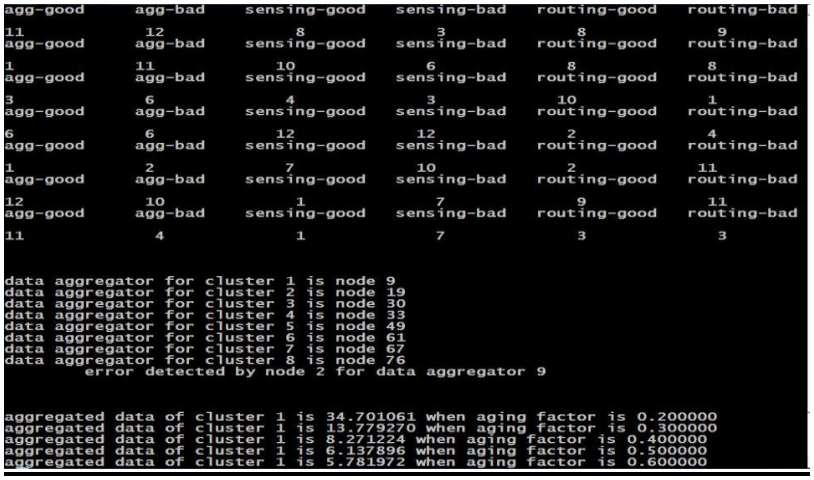

Fig –data aggregator allocation

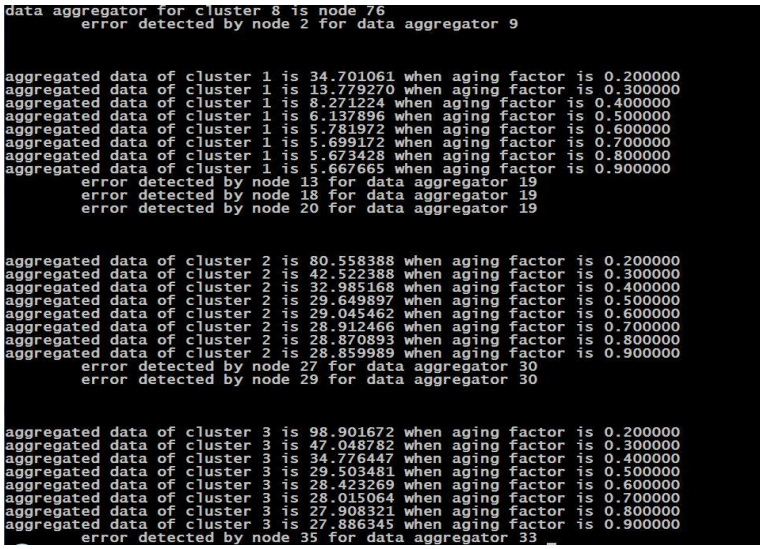

Fig – aggregated data for clusters

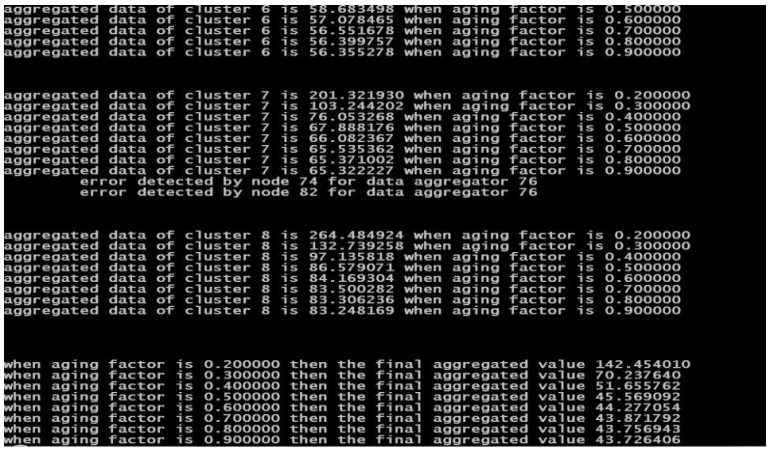

Fig - final aggregat

II. CONCLUSION

In wireless sensor networks, compromised sensor nodes can distort the integrity of aggregated data by sending false data reports and injecting false data during data aggregation. Since cryptographic solutions are not sufficient to prevent these attacks, general reputation based trust systems are proposed in the literature. This paper has presented a novel reliable data aggregation and transmission protocol (RDAT) that introduces functional reputation concept. In comparison with general reputation, the simulation results show that protocol RDAT improves the security and reliability of the aggregated data by using functional reputation concept.Future work includes the simulation of this protocol on any simulation software such as Network Simulator (NS-2) or QUALNET or any other simulation software and get the exact results and compare these results with the implementation results of another reliable data aggregation protocol i.e "Ant Colony data gathering protocol". Carrying out more detailed simulator runs would also allow the protocols to be evaluated in more detail

REFERENCES

[1] Suat Ozdemir ," Functional Reputation Based Data Aggregation for Wireless Sensor Networks", IEEE International Conference on Wireless & Mobile Computing, Networking & Communication,2008

[2] Tamara Pazynyuk, JiangZhong Li, George S. Oreku," Reliable Data Aggregation Protocol for Wireless Sensor Networks",IEEE 2008

[3] R. Rajagopalan and P.K. Varshney, "Data aggregation techniques in sensor networks: A survey", IEEE Communications Surveys and Tutorials, vol.8, no. 4, 4th Quarter 2006.

[4] Hong Luo, Qi Li, Wei Guo , "RDA: Data Aggregation Protocol for WSNs", Beijing Key Laboratory of Intelligent Telecommunications Software and Multimedia, IEEE2006 (p 1-4)

[5] Sang-ryul Shin, Jong-il Lee, Jang-woon Baek, Dae-wha Seo," Reliable Data Aggregation Protocol for Ad-hoc Sensor Network Environments", IEEE 2006

[6] Yi Yang, Xinran Wang, Sencun Zhu, and Guohong Cao, "SDAP: A Secure Hop by Hop Data Aggregation Protocol for Sensor Networks", Department of Computer Science & engineering, The Pennsylvania State University, ACM 2006 (p 1-12)

[7] H. C¸ am, S. Ozdemir, P. Nair, and D. Muthuavinashiappan, and H.O. Sanli, "Energy-Efficient and secure pattern based data aggregation for wireless sensor networks", Special Issue of Computer Communications on Sensor Networks, pp. 446-455, Feb. 2006.

[8] A. Srinivasan, J. Teitelbaum, H. Liang, J. Wu, and M. Cardei, "Reputation and Trust-Based Systems for Ad Hoc and Sensor Networks", n Algorithms and Protocols for Wireless Ad Hoc and Sensor Networks, A. Boukerche (ed.), Wiley and Sons, 2008.

[9] S. Ganeriwal and M. Srivastava, "Reputation-based framework for high integrity sensor networks", in Proc. of the 2nd ACM workshop on Security of ad hoc and sensor networks, October 2004 pp. 66-77.

[10] A. Srinivasan, J. Teitelbaum and J. Wu, "DRBTS: Distributed Reputation based Beacon Trust System", In the 2nd IEEE International Symposium on Dependable, Autonomic and Secure Computing (DASC06), Indianapolis, USA, 2006.

[11] M. Raya, P. Papadimitratos, V.D. Gligor, and J.P. Hubaux. Data-Centric Trust Establishment in Ephemeral Ad Hoc Networks, Proc. of INFOCOM, pp. 1912-1920, 2008.

[12] A. Josang and R. Ismail, The Beta Reputation System, Proc. 15th Bled Conf. Electronic Commerce, 2002.

[13] B. Przydatek, D. Song, and A. Perrig, "SIA : Secure information aggregation in sensor networks", Proc. of SenSys'03, Nov 5-7, Los Angeles,CA, 2003.

[14] K. Wu, D. Dreef, B. Sun, and Y. Xiao, "Secure data aggregation without persistent cryptographic operations in wireless sensor networks", Ad Hoc Networks, vol. 5, no.1, pp. 100-111, 2007