SOCIETY FOR SCIENCE AND EDUCATION
UNITED KINGDOM

TNC TRANSACTIONS ON NETWORKS AND COMMUNICATIONS

VOLUME 6, NO. 5
ISSN: 2054 -7420

Enhanced Intelligent Model for NCD in Wireless Sensor Networks

Manyam Thaile, O.B.V. Ramanaiah
Department of Computer Science & Engineering
JNTUH-College of Engineering, Hyderabad
manyamthaile@gmail.com, obvramanaiah@gmail.com

ABSTRACT

The serious security threat for Wireless Sensor Network comes from compromised nodes. Node Compromise (NC) effects of two types: independent and dependent. In independent type, the NC effect is limited to that node only; whereas in dependent, it will spread to all the nodes across the network. In literature, there is an intelligent model which predicts the spread of node compromise. This model suffers from false positives as it trusts the communication from neighbouring nodes. To address this issue, our Parameter Grouping (PG)†mechanism is used in association with the existing Intelligent Model (IM). This Extended Intelligent Model (E-IM) performs better than the IM. The E-IM is studied through NS-2 based simulation and it's performance is analyzed.

Keywords: WSN security; Uniform Model;Gradient Model; NCD;

1 Introduction

Wireless Sensor Networks (WSN) are used in different areas. A WSN is a self-configuring network which consists of a large number of sensor nodes and are scattered either in regular or random manner. WSNs measure environmental conditions like temperature, wind, sound, pollution levels, and so on. Due to the unattended nature of a WSN, it becomes vulnerable as an attacker can physically capture nodes to make them compromised.

Usually, the attacks are of two types: outsider and insider attacks. Outsider attacks find no extraordinary access of the deployed sensor network yet wants to harm the network. These are also known as external attacks [1]. The attacker nodes which participate and execute this type of attack are not the part of network but still authorize themselves to harm the network. In insider attacks nodes situated in the network are compromised. These attacks from inside are generated by the network nodes rather than from outside nodes, and they are truly a part of the sensor network. These types of attacks are more dangerous than that of outside attacks as the insider knows sensitive information, and has all types of access rights. Hence, to detect compromise nodes is of paramount importance in WSN security.

Compromise node effects are of two types: Independent and Dependent. In case of Independent type, the compromised node does not effect its neighbours. A lot of research work is available on independent

DOI: 10.14738/tnc.65.5396
Publication Date: 31st October, 2018
URL: http://dx.doi.org/10.14738/tnc.5396

node compromise detection (NCD). In case of dependent type, the compromise node effects its neighbours and then it spreads across the network. Little work exists in literature for dependent NCD.

The work in [2] represents an intelligent model for dependent NCD. This model estimates the compromise probability of a node based on its compromised neighbour nodes. This model suffers from the risk of false positives. This model is augmented with our Parameter Grouping (PG) model to mitigate false positivies.

The rest of the paper is organized as follows: The related work is discussed in Section-II. The network and attacker models are explained in Section-III. The Proposed Extended-Intelligent Model is presented in Section-IV, and Section-V concludes the paper.

2 Related Work

In independent NCD, behaviour of a node/zone/network is analyzed with the help of different parameters such as packet arrival rate, packet sending rate, packet arrival time, node energy, and node location [4, 5, 6]. The paper [7] introduced a Reputation-Based trust management scheme in which a Bayesian formulation is used to compute an individual node's trustworthiness. The trustworthiness evaluation frame work proposed in [8] with the help of probability and entropy concepts. The compromised nodes are usually not revoked by these schemes due to the likelihood of false positive reports.

Software attestation is another method to detect Independent NCD. This method checks integrity of software code of the sensor node. The papers [9, 10, 11, 12] present work related to software attestation method. All these schemes require each and every sensor node in that network to be attested, whereas in real time scenario all the nodes may not be compromised. Thereby, benign nodes become part of the attestation unnecessarily. This results in wastage of resources of the benign nodes. The works in papers [13, 3] have combined Reputation-based and Attestation schemes. They follow a two-step procedure:

The first step is Identifying an untrustworthy zone/ node, and the next step is Software attestation of that zone/node.

2.1 Intelligent Model

For dependent NCD, a different approach named Intelligent Model (IM) is reported in [2]. The probability of a node for compromise is estimated based on compromised neighbour nodes. This model is further classified into Uniform and Gradient.

- Uniform: Each node has the equal probability for node compromise irrespective of the node position.
- Gradient: Each node has the different probability for node compromise. The far away nodes from the Base Station (BS) have more probability for node compromise.

An intelligent means use current compromised node probability then estimate future compromise node probability.

1) Intelligent uniform model: In this model, a compromised node will have its impact on its neighbour for compromise. But the position of the node is immaterial. The compromise probability of a node is estimated with the help of all the compromised neighbour nodes of it [2]. The mathematical model can be expressed with the following:

$$NCP_k = 1 - \prod_{i=1}^{N} \prod_{j=1}^{M_i} (1 - \frac{P_i}{NG_{ij} - C_{ij}}) \quad (1)$$

- *NCPk* =Node Compromise Probability of *kth* node positioned at (*x, y*) (position is immaterial)
- N=Maximum number of hops of the node being considered
- *Mi*=Number of compromised nodes in *ith* hop of the node being considered
- *Pi*=probability of compromise of the *ith* node using PG model
- NGij =Total number of 1-hop neighbours of jth compromised node in ith hop
- Cij =Total number of 1-hop COMPROMISED neighbours of jth compromised node in ith hop

2) Intelligent gradient model: This model can be adapted in application environments namely, military where the probability of node compromise will be more if the it is far away from BS. The mathematical model is given by:

$$NCP_{(x,y,k)} = (S_k)(1 - \prod_{i=1}^{N} \prod_{j=1}^{M_i} (1 - \frac{P_i}{NG_{ij} - C_{ij}})) \quad (2)$$

A node gradient distance is calculated from BS by using the Euclidean Distance. If a node has more than specified distance from the BS then we conclude that the node is far away from the BS. It is defined as:

$$S_k = \begin{cases} 1, if(d(BS,k) > Th) \\ 0, otherwise. \end{cases} \quad (3)$$

Where d(BS, k) is distance between BS and k nodes. The localization techniques will be used to find out the position of sensor nodes [14,15]. These models generate more positives that are false and not detect false negatives effectively.

3 Network and Attacker Models

In this section, we explained our model of wireless sensor networks and the attacker model under which we observe our models.

3.1 Network Model

We assume a static WSN in which the sensor nodes do not change their positions after deployment. We also assume that all direct communication links between the sensor nodes are bidirectional. We also assume that the Base Station (BS) is a trusted entity. If the BS is compromised, the entire mission of the sensor network can be easily undermined. We assume that every sensor node is able to obtain its location information and identify its placement zone by using an existing secure localization scheme such as [14], [15]. We considered two types of networks: Flat and Hierarchical (Zone). The flat network consists of all sensor nodes have same capabilities. The sensor nodes sense environment features and process them and report to BS through multi-hop communication. The Hierarchical network is divided into number of zones. Each zone has some of the normal nodes and then from that one of the node selected as Zone Head (ZH). ZH collects all information from it's zone members and then send to the BS.

3.2 Attacker Model

An attacker is captured physically sensor nodes with the purpose to steal secret data stored. The attacker may inject malicious code onto the nodes and make them as compromised. An attacker redeploys the nodes back them into the network to launch further attacks such as Routing Attacks (e.g., Gray Hole, Black Hole, Sybil, Wormhole and Hello Flood), Identity Replication, False Data Injection and Passive Data Gathering. It is necessary to detect compromise nodes are very important.

4 Extended intelligent NCD model

We extended IM with our PG model to reduce false positives and detect false negatives effectively. An E-IM model is further classified into Extended-Intelligent Uniform (EIU) NCD model and Extended-Intelligent Gradient (EIG) NCD model. These two models implemented on flat network. We also extended E-IM model to Hierarchical network (zone).

The IM model estimates compromise probaility for a node and it uses only compromised neighbour nodes information. The compromised neighbour nodes may not give correct report about it's neighbour nodes sometimes.

The E-IM model estimates compromise probability of a node with combination of two things: Based on compromised neighbour nodes and behaviour of the node (PG).

The behaviour of the node is evaluated with the help of five parameters [3]. They are Packet sending rate, Depletion of node energy, Node location, False information, and Non-availability. These parameters generates binary values (0,1). The five parameters generate total of 32 (25 = 32) binary patterns, which are five bit in length. We considered these binary patterns as node compromise probability. The node compromise probabilities are 0, 0.2, 0.4, 0.6, 0.8, 1 as shown in the Table-I.

Table 1: Probabilities of compromise node

Pattern	1^ts	0^ts	Probability
00000	0	5	0
00001	1	4	0.2
00011	2	3	0.4
00111	3	2	0.6
01111	1	4	0.8
11111	5	0	1

4.1 Flat Networks

1) Extended Intelligent Uniform NCD Model: This model accommodate the application environment where the attackers attacking with same probability. The following mathematical model uses to estimate node compromise probability P_k:

$$P_k = \frac{P_k + NCP_k}{2} \tag{4}$$

Where

- P_k= k^{th} node compromise probability
- P_k=current compromise probability of k^{th} node using PG
- NCP_k=compromise probability of k^{th} node with respect to all compromised neighbour nodes.

To describe the EU model clearly, we use Fig. 1 in order to calculate compromise probability of the node 'j'. In Fig. 1, nodes e, d, c, i, o, p and k are 1-hop neighbors of node 'j'; nodes d, c, b, h, n, o, p and j are 1-hop neighbors of node 'i'; nodes i, d, e are compromised nodes. In Fig. 1, for node j, N=4, i,e., node j can reach all the sensors in the network within 4 hops, node j has one 1-hop neighbor node (node i), one 2-hops neighbor node (node l) which have been recently compromised. So that M_1=1, M_2=1. Node i has eight 1-hop neighbors, thus n_{11}=8. Node i has one 1 hop compromised neighbor, i.e., node d, then k_{11}=1. Node l has five 2-hops neighbors (nodes d, p, j, v, and w) and one 2-hops compromised neighbor (node d), consequently n_{21}=5, k_{21}=1. Suppose p_1=0.8 p_2=0.4, and P_j=0.2. We calculate the compromise probability of the node j's as follows:

$$P_{(x,y,j)} = \frac{0.2+1-(1-\frac{1}{8-1}*0.8)(1-\frac{1}{5-1}*0.4)}{2} = 0.20$$

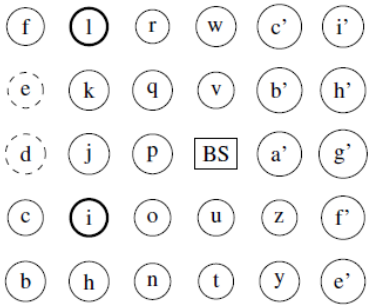

Fig. 1: Extended Intelligent Uniform Model

2) Extended Intelligent Gradient NCD Model: The given mathematical model used to estimate compromise node probability which is far away from BS.

$$P_{(x,y,k)} = \frac{P_k + S_{(x,y,k)} * NCP_{(x,y,k)}}{2} \quad (5)$$

where $S_{(x,y,k)}$= k^{th} node is far away from Base Station (1 or 0) and P_k is current compromised probability.

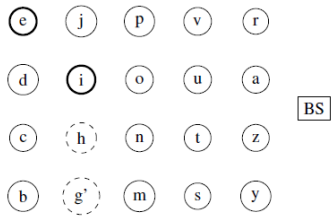

Fig. 2: Extended Intelligent Gradient Model

To describe the EG model clearly, we use Fig. 2 in order to calculate compromise probability of node 'c'. In Fig. 2, compromised nodes will not be recovered after they detected as compromised nodes. In Fig. 2, node i and e are recently compromised nodes that have been compromised in the last time period; Nodes h, g' are compromised nodes previously. In Fig. 2, for node 'c', N=4, i,e., node 'c' can reach all the sensor nodes in the network within 4 hops, node 'c' has one 1-hop neighbor node (node i), one 2-hops neighbor node (node e), that have been recently compromised. So that M1=1, M2=1, M3=0 and M4=0. Node i has eight 1-hop neighbors, thus n11=8. Node i has two one hop compromised neighbor, i.e., node e and node h, then k11=2. Node e has five 2-hops neighbors (nodes p, o, n, h, and c) and one 2-hops compromised neighbor (node h), consequently n21=5, k21=1. Suppose p1=0.8, p2=0.6, p3=0, p4=0, S(x,y,c)=1 and Pi=0.8. We calculate compromise probability of node c's as follows:

$$P_{(x,y,c)} = \frac{0.8+1*[1-(1-\frac{1}{8-2}*0.8)(1-\frac{1}{5-1}*0.6)]}{2} = 0.53$$

4.2 Hierarchical Network

We estimated compromise probability for entire network and zone-wise, for which Extended Intelligent Uniform model uses only. The network is divided into 'N' number of zones in which a node acts as Zone Head (ZH). Every ZH can measure compromise probability of it's zone. The Base Station (BS) measure compromise probability for the total network. In Fig.3, zh1, zh2, zh3 and zh4 are ZHs, remaining nodes are normal nodes. The dashed circle (node:g) is previously compromised node and thick circle (node:a) is currently compromised node. The compromise probability of total network and zone-wise is tabulated in Table-IV and V respectively.

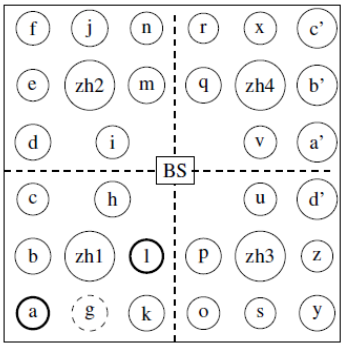

Fig. 3: Zone Network

5 Simulation Study

In this section, we describe our simulation experimental environment and then discuss the simulation results.

5.1 Simulation Environment

We considered NS2 open source simulator for the analysis of the E-IM models. We simulated flat and hierarchical (zones) networks. The network size is scaled from 20, 40, 60, 80, and 100 nodes.

5.2 Simulation Results

The performance of the E-IM models (EU, EG) are evaluated. The simulation results are compared with IM models. We considered following performance metrics.

- False Positive (FP): A trusted node is identified as untrusted.
- False Negative (FN): An untrusted node is identified as trusted.
- False Positive Ratio (FPR): $\frac{FP}{FP+FN}$.
- False Negative Ratio (FNR): $\frac{FN}{FP+FN}$.

Table 2: Flat network

No.of Nodes	IM		E-IM		IM		E-IM	
	False Positives	False Negatives	False Positives	False Negatives	FPR	FNR	FPR	FNR
20	4	1	1	2	0.80	0.20	0.33	0.66
40	8	2	3	4	0.80	0.20	0.42	0.57
60	14	3	7	6	0.82	0.17	0.53	0.46
80	19	4	10	8	0.82	0.17	0.55	0.44
100	26	5	14	9	0.83	0.16	0.59	0.40

The Table-II shows results of false positives, false negatives, false positive ratio, and false negative ratio. We tested three cases for each network of size 20, 40, 60, 80, and 100 nodes. The E-IM model reports less false positives when compare with IM model. An E-IM model is detected false negatives effectively when compare with IM model.

The Fig.4 shows false positive rates between IM and E-IM models. The x-axis indicates the number of nodes, and y-axis indicates false positive rates. An E-IM model got less false positives when compare with IM model.

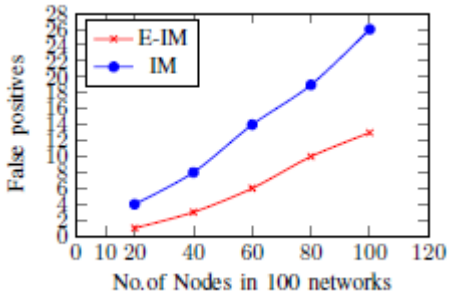

Fig. 4: False positives

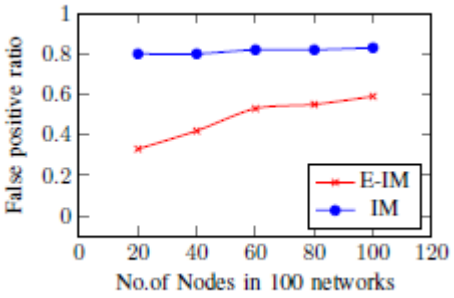

Fig. 5: False positive ratio

The Fig.5 shows the false positive ratio between IM model and E-IM model. The x-axis indicates the number of nodes, and y-axis indicates false positive ratio. The figures shows an E-IM model got less false positive ratio when compare with IM model.

The Fig.6 and 7 show the false negative rates and false negative ratio respectively. Fig.6 represents an E-IM model detected false negatives effectively.

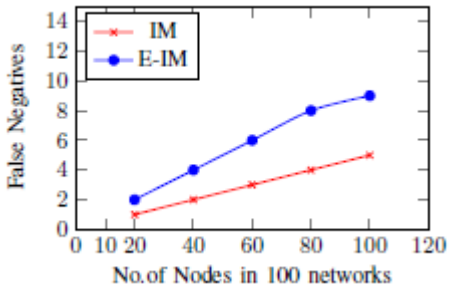

Fig. 6: False Negatives

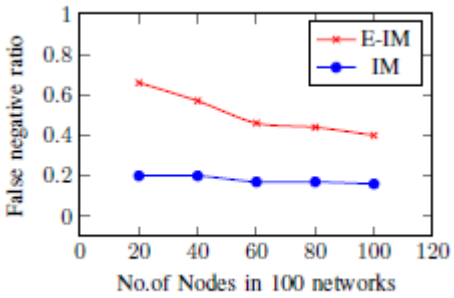

Fig. 7: False negative ratio

We calculated node probabilities for EU and EG. The IM model has given high probabilities but it leads to more false positivies. The E-IM provides less probabilities when compare with IM model and it gives less false positives. The values are shown in Table-III.

Table 3: Node probability

No.of Nodes	Uniform		Gradient	
	IM	E-IM	IM	E-IM
20	0.10	0.05	0.19	0.10
40	0.26	0.12	0.25	0.12
60	0.42	0.31	0.31	0.15
80	0.60	0.36	0.35	0.21
100	0.73	0.52	0.40	0.30

The Fig.8, 9 shows node compromise probability between IM model and E-IM model. An E-IM model gives less compromise probability.

Table-IV indicates the entire network compromise probability. We can predict whether network is going to be compromised or not in the future. We tested three cases for finding network probability by using zones probability. In the table $z_1 = 6 - 1$

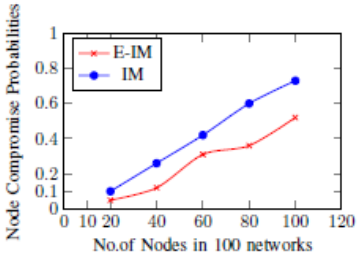

Fig. 8: Uniform model

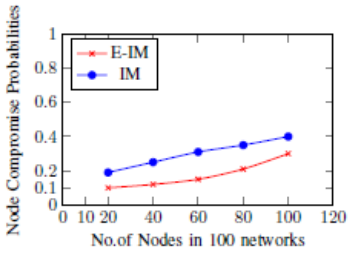

Fig. 9: Gradient model

indicates the difference between total number of nodes (6) in zone z_1 and number of compromised nodes (1) in zone z_1.

Table 4: Network probability by using zones in hierarchical networks

No.of Nodes	No.of Zones	No.of Nodes in each zone and compromised nodes			Probability			Average
		Case1	Case2	Case3	Case1	Case2	Case3	
20	2	z_1=6-1,z_2=10-1	z_1=6-1,z_2=9-1	z_1=6-1,z_2=10-1	0.06	0.06	0.06	0.06
40	4	z_1=9-1,z_2=8-1, z_3=10-2,z_4=7-0	z_1=9-1,z_2=19-2, z_3=2-0,z_4=3-1	z_1=11-2,z_2=6-1, z_3=5-0,z_4=11-1	0.11	0.15	0.12	0.12
60	6	z_1=10-2,z_2=14-2, z_3=9-1,z_4=3-0, z_5=8-1,z_6=4-0	z_1=1-0,z_2=4-0, z_3=5-1,z_4=11-2, z_5=14-2,z_6=5-1	z_1=7-0,z_2=10-2, z_3=3-0,z_4=11-2, z_5=7-1,z_6=10-1	0.15	0.18	0.16	0.16
80	8	z_1=3-0,z_2=8-1, z_3=17-3,z_4=4-0, z_5=6-1,z_6=14-2, z_7=5-0,z_8=6-1	z_1=14-2,z_2=5-1, z_3=17-2,z_4=6-1, z_5=9-1,z_6=8-1, z_7=3-0,z_8=6-0	z_1=6-1,z_2=6-1, z_3=1-0,z_4=9-1, z_5=10-1,z_6=12-2, z_7=11-2,z_8=3-0	0.19	0.21	0.21	0.20
100	10	z_1=11-2,z_2=11-1, z_3=4-0,z_4=3-0, z_5=6-0,z_6=14-2, z_7=10-1,z_8=17-4, z_9=6-0,z_{10}=3-0	z_1=1-0,z_2=9-1, z_3=10-1,z_4=6-1, z_5=4-0,z_6=4-0, z_7=10-1,z_8=5-1, z_9=11-2,z_{10}=16-3	z_1=6-0,z_2=12-2, z_3=9-1,z_4=8-1, z_5=9-1,z_6=4-0, z_7=3-0,z_8=22-4, z_9=11-1,z_{10}=1-0	0.20	0.31	0.19	0.23

Table 5: Zone-Wise probability in hierarchical net- works

No.Of Nodes	No.Of Zones	Zone-Wise Probability
20	2	z_1=0.96, z_2=0.97
40	4	z_1=0.96,z_2=0.96,z_3=0.31, z_4=0.62
60	6	z_1=0.31,z_2=0.63,z_3=0.64, z_4=0.63,z_5=0.96,z_6=0.64
80	8	z_1=0.64,z_2=0.96,z_3=0.64, z_4=0.64,z_5=0.96,z_6=0.94, z_7=0.31, z_8=0.32
100	10	z_1=0.31,z_2=0.97,z_3=0.64, z_4=0.64,z_5=0.32,z_6=0.32, z_7=0.64,z_8=0.94,z_9=0.64, z_{10}=0.31

The Fig.10 shows network compromise probability and it can be observed that when compromised nodes are increased in zones then compromised probability is also increased.

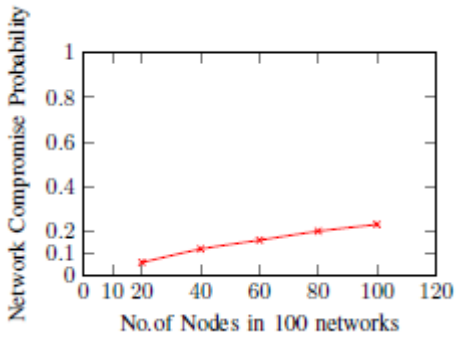

Fig. 10: Network compromise probability

We measured compromise probability zone-wise with the help of number of nodes in a zone and also number of compromised nodes. If any zone consist of compromised nodes then that zone has more compromised probability as shown in Table-V.

6 Conclusion

We implemented an E-IM models to measure compromise probability of a node in flat network. We also extended an E- IM model to hierarchical network and analyzed compromised probability for zone-wise and network wise. The E-IM model reduced false positivies and detected false negatives effectively. We compared performance of E-IM model with IM model and got better performance.

REFERENCES

[1] Kalpana Sharma and M K Ghose, ”Wireless Sensor Networks: An Overview on its Security Threats”, IJCA Special Issue on Mobile Ad-hoc Networks (MANETs), Volume 1, Issue 8, pp.42-45, 2010.

[2] Xiangqian Chen, Kia Makki, Kang Yen, Niki Pissinou, ”Node Compromise Modeling and its Applications in Sensor Networks,” in the proceedings of IEEE ISCC 2007, IEEE Symposium on Computers and Communications, Aveiro, Portugal, July 2007.

[3] Manyam Thaile, and O.B.V Ramanaiah, ”Node Compromise Detection Based on Parameter Grouping in Wireless Sensor Networks,” SECURWARE 2016:The Tenth International Conference on Emerging Security Information, Systems and Technologies, July. 24 – 28, 2016, pp. 14-20, ISBN: 978-1- 61208-493-0, Nice, France.

[4] F. Li and J. Wu, “Mobility Reduces Uncertainty in MANET,” May 2007, Proc. IEEE International Conference on Computer Communications, pp. 1946-1954.

[5] D. Estrin, R. Govindan, J. Heidemann, S. Kumar, ”Next century challenges: scalable coordination in sensor networks,” ACM MobiCom’99, Washingtion, USA, 1999, pp. 263-270.

[6] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, ”A survey on sensor networks,” WEEE communications Magazine, Volume: 40 Issue: 8, pp. 102-114, August 2002.

[7] S. Ganeriwal and M. Srivastava, ”Reputation-Based Framework for High Integrity Sensor Networks,” Oct. 2004, Proc. ACM Workshop Security of Ad Hoc and Sensor Networks (SASN), pp. 66-77.

[8] Y. Sun, Z. Han, W. Yu, and K. Liu, ”A Trust Evaluation Framework in Distributed Networks: Vulnerability Analysis and Defense against Attacks,” Apr. 2006, Proc. IEEE INFOCOM, pp. 1-13.

[9] A. Seshadri, A. Perrig, L. van Doorn, and P. Khosla, ”SWATT:SoftWare-Based Attestation for Embedded Devices,” Proc. IEEE Symp. Security and Privacy, May 2004.

[10] T. Abuhmed, N. Nyamaa, and D. Nyang, ”Software-Based Remote Code Attestation in Wireless Sensor Network,” Proc. of IEEE GLOBECOM, December. 2009.

[11] T. Park and K. G. Shin, ”Soft Tamper-Proofing via Program Integrity Verification in Wireless Sensor Networks,” IEEE Trans.Mobile Computing, May/June 2005, vol. 4, no. 3, pp. 297-309.

[12] Y. Yang, X. Wang, S. Zhu, and G. Cao, ”Distributed Software-Based Attestation for Node Compromise Detection in Sensor Networks,” Proc.IEEE 26th Int’l Symp. Reliable Distributed Systems (SRDS), Oct. 2007.

[13] Jun-Won Ho, Matthew Wright, and Sajal K. Das, ”ZoneTrust: Fast Zone-Based Node Compromise Detection and Revocation in Wireless Sensor Networks Using Sequential Hypothesis Testing,” IEEE Transactions on Dependable and Secure Computing, July/August 2012, vol. 9, no. 4, pp. 494-511.

[14] S. Capkun and J.P. Hubaux, ”Secure Positioning in Wireless Networks,” IEEE J. Selected Areas in Comm., vol. 24, no. 2, pp. 221-232, Feb. 2006. [15] Z. Li, W. Trappe, Y. Zhang, and B. Nath, ”Robust Statistical Methods for Securing Wireless Localization in Sensor Networks,” Proc. Fourth Int’l Symp. Information Processing in Sensor Networks (IPSN), Apr. 2005.