Dempster-Shafer Theory to Identify Insider Attacker in Wireless Sensor Network

Muhammad Ahmed, Xu Huang, Dharmendra Sharma

▶ **To cite this version:**

Muhammad Ahmed, Xu Huang, Dharmendra Sharma. Dempster-Shafer Theory to Identify Insider Attacker in Wireless Sensor Network. 9th International Conference on Network and Parallel Computing (NPC), Sep 2012, Gwangju, South Korea. pp.94-100, 10.1007/978-3-642-35606-3_11 . hal-01551336

HAL Id: hal-01551336

https://inria.hal.science/hal-01551336v1

Submitted on 30 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Dempster-Shafer Theory to Identify Insider Attacker in Wireless Sensor Network

Muhammad Ahmed[1] , Xu Huang[1], and Dharmendra Sharma[1],

[1] Faculty of Information Sciences and Engineering,
University of Canberra, Australia
{muhammad.ahmed, xu.huang,dharmendra.sharma}@canberra.edu.au

Abstract. Due to the construction and network infrastructure of wireless sensor network (WSN) are known to be vulnerable to variety of attacks. In order to ensure its functionality especially in malicious environments, security mechanisms are essential. Several works have been done to secure WSN, but identification of insider attacker has not been given much attention. In the WSN system the malicious node behavior is different from the neighbor nodes. Instead of relying the untrustworthy neighbor node we use Dempster-Shafer theory (DST) of combined evidence to identify the insider attacker in WSN. This theory reflects with the uncertain event or uncertainty as well as uncertainty of the observation. The mathematical calculation shows the DST capability of identifying the insider attacker.

Keywords: *Wireless Sensor Networks, Insider Attacker, Security, Dempster-shafer theory*

1 Introduction

Wireless sensor networks are a new technology for collecting data with autonomous sensors. Recently, this technology became more popular because of its application and cost. It consists of large number of low cost, low power and multifunctional sensors embedded with short range wireless communication capability. Sink in which all data is transmitted in an autonomous way has high capacity of storage and analysis power. The application of WSN includes battlefield surveillance, border monitoring, habitat monitoring, intelligent agriculture, home automation, etc.

In this information age the world is interconnected via various communications. Security provisioning is a critical requirement for any communication network. Security in the wireless sensor network is challenging and important task because of its characteristics that include, open nature of wireless medium, unattended operation, limited energy, memory, computing power, communication bandwidth, and communication range. Considering those characteristics many algorithms have developed for the secure functionality of WSN. Most of the work has focused on the pair wise key establishment, authentication access control and defense against attack.

Most importantly those works mainly focused on the traditional cryptographic information, data authentication in order to build the relationship between the sensors. However, the unreliable communications through wireless channel made the communication technique vulnerable by allowing the sensor nodes to compromise and release the security information to the adversary [1]. The compromised entity of the network acts as a legitimate node. So it is easy for the adversary to perform the insider attacks. When insider attack occurs for a node, this node will behave abnormally such as tampering the massage from other member, dropping the data or broadcast excessive data.

So far, not much attention has been given to save the network from the insider attacker that caused by the abnormally behaved node. In this paper, we have proposed Dempster–Shafer theory (DST) based insider attacker identification mechanism with neighbor nodes parameters observation as DST has the feature of dealing with uncertainty. In our proposed method the system does not need to have any prior knowledge of the pre-classified training data of the nodes.

The paper is organised as follows: section 2 is comprised of the overview of the related work followed by the system architecture and network model in section 3. The detail of the dempester-shafer theory for insider attacker identification process is described in section 4. The evaluation in WSN and mathematical calculation is given in section 5 followed by conclusion in section 6.

2 Related Work

To identify insider attacker in wireless sensors networks several work has been done in the past but DST based method was not given significant attention.

For detection of abnormal behavior of the nodes or insider attacker Staddon et al [2] proposed to trace the failed nodes in sensor networks at the base station assuming that all the sensor measurement will be directed along the sinker based on the routing tree. In this work the sinker has the global view of the network topology and can identify the failed nodes through route update message and it is directional.

Watchdog like technique was proposed by Marti [3], this technique can detect the packet dropping attack by letting nodes listen to the next hope nodes broadcasting transmission. In this multiple watchdog work collaboratively in decision making and reputation system is necessary to provide the quality rating of the participants.

Zhang et al [4] proposed a scheme which is the first work on intrusion detection in wireless ad hoc networks. A new architecture is investigated for collaborative statistical anomaly detection which provides protection from attack on ad hoc routing.

These developments somehow solve the mathematical problems with certain constrain but does not take the insider attacker identification in consideration with the uncertainty of observation by neighbor nodes.

3 System Architecture and Network Model

In our system we have considered the neighbor nodes or observer nodes evidence to identify the insider attacker. The neighbor nodes will share their independent observation about the suspected insider attacker behavior. The data from the neighbor nodes we will consider as evidence, which can be in the form of malcounts (number of occurrences of misbehavior). We will combine the independent pieces evidence and take the decision based on the DST.

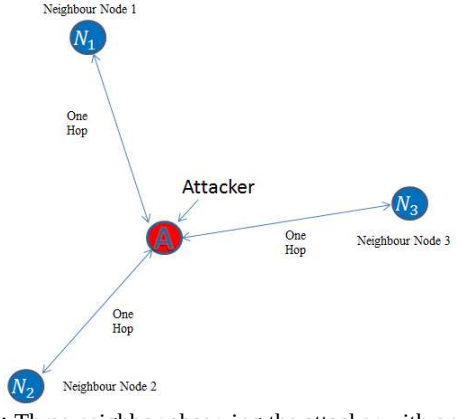

Fig 1: Three neighbor observing the attacker with one hop

In WSN the neighbor with one hop will observe the data as node behavior. Temperature measurement wireless sensor network scenario neighbor will check the temperature reading and that will be become the evidence. The neighbors can obtain degrees of belief about the proposition from related proposition subjective probabilities. In the figure (1), neighbor nodes N_1, N_2 and N_3 will share their independent observation about the insider attacker before taking the decision. The neighbor nodes will be the nodes with nearest euclidian distance.

4 Methodology

The Bayesian theory is the canonical method for statistical inference problems. The Dempster-Shafer decision theory is considered a generalized Bayesian theory. It allows distributing support for proposition, not only to a proposition itself but also to the union of propositions that include it. [5] In Dempster-Shafer Theory (DST) a node can hold either supportive or uncertain opinion toward an event. It addresses the solution by representing the uncertainty in the form of belief functions. The idea is neighbor or observer nodes can obtain degree of belief about the proposition from the related proposition's subjective probabilities.

4.1 Bayesian Interface

In order to understand the Dempster-Shafer Theory beyesian approach is often studied. Bayesian inference derives a posterior probability distribution as a consequence of two antecedents, a prior probability and likelihood, probability model for the data to be observed. [6]Bayesian inference computes the posterior probability by conditioning, according to the rule of Bayes for proposition of *H* and Evidence *E*.

$$P(H)=\frac{P(E\mid H)P(H)}{P(E)} \tag{1}$$

According to Bayesians interpret, *P(H)* the priori reflects the initial degree of belief in *H* in the absence of evidence *E*. *P(H|E)*, the posteriori probability as a measure of belief about a hypothesis or proposition *H* that updates in response to evidence.

In figure one we consider node N_1, N_2 and N_3 has the representative pieces of evidence e_{N_1}, e_{N_2} and e_{N_3}, in order to support the hypothesis *H*. So, the posteriori probability becomes

$$P(H\mid e_{N_1},e_{N_2},e_{N_3})=\frac{P(e_{N_1},e_{N_2},e_{N_3}\mid H)P(H)}{P(e_{N_1},e_{N_2},e_{N_3}\mid H)P(H)+ P(e_{N_1},e_{N_2},e_{N_3}\mid\sim H)(1-P(H))} \tag{2}$$

In which ~H is not H hypothesis means node A is an attacker. The neighbor nodes observes the attacker independently, hence the computation of the equation 2 can be simplified as in equation 3 by factorization process.

$$P(H\mid e_{N_1},e_{N_2},e_{N_3}) = P(e_{N_1}H)P(e_{N_2}H)P(e_{N_3}H) \tag{3}$$

Complete knowledge of the prior and conditional probabilities is a signeficunt requirement for this approach which is difficult to determine in practice. In this approach estimation of the prior probabilities is done from the empirical data. Hence, this method does not have capability to deal with the states of ignorance.

4.2 Dempster-Shafer Framework

In DST, probability is replaced by an uncertainty interval bounded by belief and plausibility. Belief is the lower bound of the interval and represents supporting evidence. Plausibility is the upper bound of the interval and represents the non-refuting evidence. In this reasoning system, all possible mutually exclusive hypothesis (or events) of the same kind are enumerated in the frame of discernment Ω. A basic belief assignment (BBA) or mass function is a function m: $2^{\Omega}\rightarrow$ [0, 1], and it satisfies two following conditions

$$m(\partial) = 0 \tag{4}$$

$$\sum_{A \subseteq \Omega} m(A_j) = 1 \tag{5}$$

In which ∂ is the empty set and a BBA that satisfy the condition $m(\partial) = 0$. The basic probability number can be translated as $m(A)$ because the portion of total belief assigned to hypothesis A, which reflects the evidences strength of support. The assignment of belief function maps each hypothesis B to a value $bel(B)$ between 0 and 1. This defined as

$$bel(B) = \sum_{j:A_j \subseteq A} m(A_j) \tag{6}$$

The upper bound of the confidence interval is the plausibility function, which accounts for all the observations that do not rule out the given proposition. It maps each hypothesis B to a value $pls(B)$ between 0 and 1, can be defined as follows.

$$pls(B) = \sum_{j:A_j \cap B \neq \partial} m\ (A_j) \tag{7}$$

The plausibility function is a weight of evidence which is non-refuting to B. equation (8) shows the relation between belief and plausibility.

$$pls(B) = 1 - bel(\sim B) \tag{8}$$

The hypothesis not B is representing by $\sim B$. The functions basic probability numbers, belief and plausibility are in one-to-one correspondence and by knowing one of them, the other two functions could be derived.

Assuming $m_1(A)$ and $m_2(A)$ are two basic probability number by two independent items of evidence means two independent neighbor node which act as observers in the same frame of discernment. The observations (the pieces of evidence) can be combined using Dempster's rule of combination (known as orthogonal sum) as in equation (9).

$$m(B) = (m_1 \oplus m_2)(B) = \frac{\sum_{i,j:A_i \cap A_j = B} m_1(A_i)\, m_2(A_j)}{1 - \sum_{i,j:A_i \cap A_j = \partial} m_1(A_i)\, m_2(A_j)} \tag{9}$$

More than two belief function can be combined with pairwise in any order

5 Evaluation in WSN

In temperature collection WSN we consider the normal temperature range is T= 8 to 10 degree centigrade based on the Gaussian distributing with 1 sigma based on the approach taken by holder *el at* [7] , and ~T means the temperature is out of range and

consider the node *A* is attacked. So, the frame of discernment consists of two probabilities concerning the attacker node *A*: Ω = {T, ~T}. Hence, for Ω the power set has three focal elements: hypothesis *H* = {~T}, *H* = {T} and universe hypothesis *U*= Ω meaning node *A* is either attacked or a good node. We consider that neighbor node N_1 is a trusted node with the probability β. Based on the node N_1 information if node *A* is an attacker, the basic probability assignment will be as follows.

$$m_1(H) = 0;$$
$$m_1(\sim H) = \beta;$$
$$m_1(U) = 1 - \beta; \quad (10)$$

If *A* is a good node the basic probability assignment will be

$$m_1(H) = \beta;$$
$$m_1(\sim H) = 0;$$
$$m_1(U) = 1 - \beta; \quad (11)$$

Using the same approach we can construct the basic probability assignment m_1 and m_2 for neighbor node N_2 and N_3.

The combined belief of N_1, N_2 and N_3 in *H* is bel(H) = m(H) = $m_1(H) \oplus m_2(H) \oplus m_3(H)$ following the Dempster rule of combination based on equation (9). It is possible to combine any pair of arguments and then combine the remaining argument. m_1 and m_2 combination can be written as follows.

$$(m_1 \oplus m_2)(H) = \frac{1}{k}[m_1(H)m_2(H) + m_1(H)m_2(U) + m_1(U)m_2(H)]$$
$$(m_1 \oplus m_2)(\sim H) = \frac{1}{k}[m_1(\sim H)m_2(\sim H) + m_1(\sim H)m_2(U) + m_1(U)m_2(\sim H)]$$
$$(m_1 \oplus m_2)(U) = \frac{1}{k}[m_1(U)m_2(U)] \quad (12)$$

Where,

$$K = m_1(H)m_2(H) + m_1(H)m_2(U) + m_1(U)m_2(H) + m_1(\sim H)m_2(\sim H) + m_1(\sim H)m_2(U) + m_1(U)m_2(\sim H) + m_1(U)m_2(U)$$

After combining the reports from the neighbor's nodes we can identify the insider attacker.

5.1 Example

In the paper we have given some mathematical calculation and results for the combined degree of belief that the node *A* is insider attacker

Table 1: Combine degree of belief calculation

Trust probability of the neighbor node			
N_1	N_2	N_3	Combined degree of Belief
0.9	0.8	0.2	0.975
0.2	0.2	0.9	0.878
0.8	0.8	0.8	0.828

In the table 1 we can see that the calculation is done by assigning the different trust probability to the neighbor and combine degree of belief is 0.975, 0.878, 0.828 respectively. From the high belief is it concluded that the node is an attacker.

6 Conclusion

In this paper an insider identification framework in wireless sensor network is proposed with Dempster-Shafer theory of evidence combination method. the mathematical calculation shows that the result depends on the neighbor nodes reliability. Moreover, the conflict increases with the number of sources.

In future, we would like to create a database for the nodes normal behavior form that we can decide about the reliability of the nodes and employ extended dempster-shefer theory.

References

1 Zhou Y., Fang Y., and Zhang Y., "Securing wireless sensor networks: a survey," IEEE Communications Surveys & Tutorials, 3rd Quarter (2008).

2 Staddon J., Balfanz D., and Durfee G., "Efficient tracing of failed nodes in sensor networks," in WSNA 2002, pp. 122-130, Atlanta, USA (2002)

3 Marti S., Giuli T.J., Lai K., Baker M., "Mitigating Routing Misbehavior in Mobile Ad Hoc Networks," ACM MOBICOM 2000, pp. 255-265, Boston, USA, August (2000)

4 Zhang Y., Lee W., "Intrusion Detection in Wireless Ad-hoc Networks," ACM MOBICOM 2000, pp. 275-283, Boston, USA, August (2000).

5 Sentz K., "Combination of Evidence in Dempester-Shafer Theory", System Science and Engineering Department, Binghamton University, SAND 2002-0835, April (2002)

6 Koks D., Challa S., "An Introduction to Bayesian and Dempster-Shafer Data Fusion" , Published by DSTO Systems Sciences Laboratory, Australia, November (2005)

7 Holder C., Boyles R., Robinson P., Raman S., and Fishel G., "Calculating a daily Normal temperature range that reflects daily temperature variability", American Meteorological Society, June (2006)