A three-tier scheme for sybil attack detection in heterogeneous IWSN

Hui Wang*

Department of Mathematics and Computer Engineering, Ordos Institute of Technology, Ordos, China

Keywords: IWSN, Sybil attack, Three-tier, High-energy.

Abstract. Industrial wireless sensor network is a new application of wireless sensor network in the field of industry in recent years. However, IWSN may be subject to different attacks and security risks, among which the Sybil attack is the most harmful type of base attack. According to the characteristics of wireless sensor networks in industrial environment, a new three-tier detection scheme is proposed. In the first-level, all common nodes and Sybil nodes were detected by RSSI-based quadratic difference method. In the second layer, the residual energy-based method is used to continue the detection of the nodes that have been detected in the first-level. The detection of first and second-level high-energy nodes is finally completed by the base station. The simulation results show that our proposed scheme significantly improves network lifetime and effectively improve the detection rate of Sybil nodes.

1 Introduction

Wireless sensor network (WSN) is a self-organizing distributed network system, which composed of a large number of tiny sensor nodes with wireless communication and computing capabilities [1]. It collects and processes the information from the target area by cooperate with each node. Wireless sensor networks have broad application prospects in military reconnaissance, environmental monitoring, medical health monitoring, agricultural cultivation and industrial production control [2]. The WSN utilized in industrial (IWSN) is used for controlling and monitoring various industrial tasks [3]. Due to the nodes in industry wireless network are self-organizing and self-healing, they can adapt to environmental changes to perform efficiently. IWSN is generally one-time arrangement, and the possibility of later energy supplement is very small. Therefore, how to save energy and extend the survival time of the network has become the primary consideration. To solve this problem, a heterogeneous IWSN is proposed, especially IWSN based on energy heterogeneous. Although we have solved the network problems caused by the conditions of network nodes, in the actual use process, IWSN often suffers from different attacks and security threats, such as Wormhole attack, Hello flood, Sinkhole and Sybil attack. Sybil attack is one of the most harmful and easy attacks. In Sybil attack, malicious nodes are disguised as multiple nodes, for example, by imitating the target node or only by declaring false identity. Therefore, in

* Corresponding author: 923762038@qq.com

IWSN, only relying on identity authentication and encryption and other defensive measures is not enough to resist malicious attacks. The effective Sybil attack monitoring scheme is urgently needed to maintain the normal operation of the factory.

At present, various protection mechanisms have been developed to protect sensor nodes from Sybil attack. The Sybil attack was first described by Douceur in the context of peer-to peer networks [4]. In [5], it proposed that wireless sensor networks could use voting for a number of tasks. The Sybil attack could be used to "stuff the ballot box" in any such vote. However, the Sybil node can use multiple identities to vouch for each other, and it would influence the outcome of voting. In [6], it proposed the notion of location-based keys (LBKs) by binding private keys of individual nodes to both compromised nodes IDs and geographic locations. The method is computationally complex and requires many resources per node. In [7], the author argued the node identities are verified simply by analyzing the neighboring node information of each node. In [8], a lightweight solution for Sybil attack problem based on received signal strength indicator (RSSI) readings of messages is proposed.

However, previous studies did not consider how to detect in the energy heterogeneous IWSN. On the contrary, they only considered how to detect Sybil nodes in the heterogeneous network. In order to improve the network life, it is generally deployed as energy heterogeneous network in IWSN deployment.

In this paper, we propose a three-tier detection scheme for heterogeneous IWSN monitoring. Unlike previous schemes, the mechanism does not utilize the neighbor's information or location information. In our scheme, we use three kinds of nodes with different energy levels to form a network, namely normal node, first energy node and second energy node. We use the quadratic difference based on RSSI for detection. In the second layer, the secondary high energy node layer, we use the principle of residual energy to further detect. Finally, the final detection is performed by the base station.

2 Three-tier detection scheme

In our proposed scheme, Sybil node detection rate, IWSN network life and energy consumption were considered comprehensively. In this scheme, heterogeneous IWSN is composed of three sensor nodes, namely normal node, primary energy node and secondary energy node. In order to ensure that every normal node can communicate with its adjacent second-level high-energy nodes, two first-level high-energy nodes are deployed in the communication range of each normal node, and one second-level high-energy nodes is deployed in the communication range of each first-level high-energy node. Each high-energy node can balance the network cost, help the base station detect Sybil attack node and pass important information, show in figure 1. In our scheme, the second high energy node only receives control packets from the first high energy node. However, only normal nodes are eligible to be selected as cluster heads to participate in other important activities in the network.

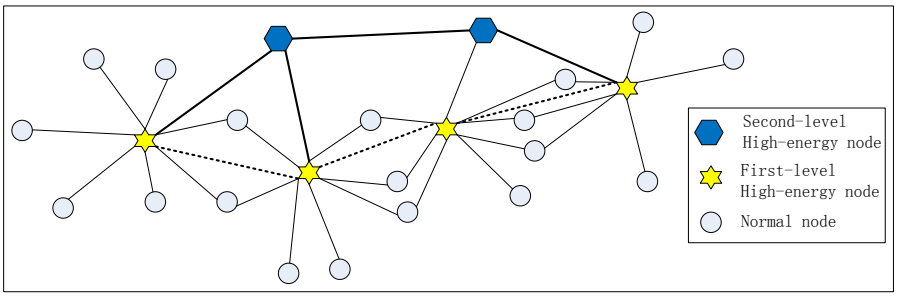

Fig. 1. The network model.

2.1 RSSI Detection based on Large-scale Fading Model

In the first layer, we use the large-scale fading model to calculate the RSSI value, thus using RSSI to detect the Sybil node.

One obvious feature of the Sybil attack is that malicious nodes forge multiple identities, which convincing the base station that the data is being transmitted from different nodes. But in fact, the forged identities are nodes in the same geographical location. A method based on RSSI is proposed to detect these malicious nodes.

In actual industrial network environment, the distance between transmitter and receiver is relatively long, so it usually includes path loss and shadow fading problems [9]. Generally, when the relative position between the receiver and the transmitter changes within the range 1~10m, the average power of the received signal basically remains unchanged, but when their relative position changes far beyond the above range, the average power of the received signal will change by several orders of magnitude. Large-scale fading is used to describe the change of the average power value of the received signal, when the distance between the receiver and the transmitter has a large-scale change. The relation between the received signal power P_r and distance d is usually expressed by Eq. 1.

$$\mathrm{P_r} \propto \frac{1}{\mathrm{d}^{\alpha}} \tag{1}$$

Here,$\alpha \geq 2$, is the path loss, can be 3~4. It is generally believed that the actual loss obeys logarithmic normal distribution. In addition to path loss, large-scale fading also includes shadow fading. Shadow fading makes the actual loss to be a random variable. Considering the factors of path loss and shadow fading, the RSSI calculation method under large-scale fading model can be obtained as Eq. 2.

$$\mathrm{R(dB)} = \mathrm{P_0(dB)} + 10\alpha\log\left(\frac{\mathrm{d}}{\mathrm{d_0}}\right) + \mathrm{X_\sigma} \tag{2}$$

Here, R is the RSSI value calculated by the receiver, P_0 is the transmitting power, α is the path loss, d represents the distance between the receiving node and the transmitting node, d_0 is the reference distance, which is *1m* indoors and *100m* or *1km* outdoors. X_σ represents path loss due to shadow fading, $X_\sigma \sim N(0, \sigma^2)$. In the most empirical equations, the standard deviation σ^2 can be *4~12*dB.

Assume that at time *ti*, node N_1 transmits control packets to the two adjacent first-level high-energy nodes *Fh1* and *Fh2*. The first-level high-energy nodes receive control packets from N_1, and use equation 2 to calculate its RSSI value, and get Eq. 3.

$$\mathrm{R_{i1}(dB)} = \mathrm{P_0(dB)} + 10\alpha\log\left(\frac{\mathrm{d_{Fh1}}}{\mathrm{d_0}}\right) + \mathrm{X_{i\sigma 1}} \tag{3}$$

Here, P_0 is the transmitting power, α is the path loss, d_{Fh1} represents the distance between the N_1 and *Fh1*, d_0 is the reference distance, $X_{\sigma 1} \sim N(0, \sigma^2)$ is the shadow fading.

For Fh2, another first-level high-energy node, which receives the control packet from N_1 node, we can also calculate its RSSI value, as shown in Eq. 4.

$$\mathrm{R}_{i2}(dB) = P_0(dB) + 10\alpha log\left(\frac{di_{Fh2}}{d_0}\right) + X_{i\sigma 2} \tag{4}$$

Here, d_{Fh2} represents the distance between the N_1 and *Fh2*, $X_{\sigma 1} \sim N(0, \sigma^2)$, the other parameters are the same as above.

According to equation3 and 4, the RSSI D-value between *Fh1* and *Fh2* can be calculated in Eq. 5.

$$\Delta \mathrm{R}_i = R_{i1}(dB) - R_{i2}(dB) = 10\alpha log\left(\frac{di_{Fh1}}{d_{iFh2}}\right) + \Delta X_{i\sigma} \quad (5)$$

Here, $\Delta X_{i\sigma} = X_{i\sigma 2} - X_{i\sigma 1}, \Delta X_{i\sigma} \sim N(0,2\sigma^2)$.

Assume at time t_j, node N_2 transmits control packets to the two adjacent first-level high-energy nodes *Fh1* and *Fh2*. Similarly, node N_2 also can obtain the RSSI D-value from the first-level high-energy nodes *Fh1* and *Fh2*, as shown in Eq. 6.

$$\Delta \mathrm{R}_j = R_{j1}(dB) - R_{j2}(dB) = 10\alpha log\left(\frac{d_{jFh1}}{d_{jFh2}}\right) + \Delta X_{j\sigma} \quad (6)$$

Using equation 5 and 6, we can calculate the RSSI quadratic difference value, as shown in equation 7.

$$\Delta \mathrm{R} = \Delta R_i - \Delta R_j = 10\alpha log\left(\frac{d_{iFh1}d_{jFh2}}{d_{iFh2}d_{jFh1}}\right) + \Delta X_{ij} \quad (7)$$

Here, $\Delta X_{ij} = \Delta X_{i\sigma} - \Delta X_{j\sigma}, \Delta \mathrm{X}_{ij} \sim N(0,4\sigma^2)$.

If node N_1 is a malicious node, node N_2 is its forged Sybil node, according to the Sybil attack principle, node N_1 and node N_2 are actually in the same physical location. Therefore, $d_{iFh1} = d_{jFh1}, d_{iFh2} = d_{jFh2}, \frac{d_{iFh1}d_{jFh2}}{d_{iFh2}d_{jFh1}} = 1$, thus $10\alpha log\left(\frac{d_{iFh1}d_{jFh2}}{d_{iFh2}d_{jFh1}}\right) = 0$, let it substitute into Eq. 7,we can obtain Eq. 8.

$$\Delta \mathrm{R} = \Delta X_{ij} \quad (8)$$

Here, $\Delta \mathrm{R} \sim N(0,4\sigma^2)$. The probability density function of ΔR can called as Eq.9.

$$\mathrm{f}(\Delta R) = \frac{1}{2\sqrt{2\pi}\Delta R} e^{-\frac{\Delta R^2}{8\Delta R^2}} \quad (9)$$

As a general rule, the confidence level is usually 95%. When ΔR falls in this range, we think there is a Sybil attack.

2.2 Residual energy-based sybil attack detection

In industrial production, in order to ensure the quality and quantity of automated production threads, the detection rate of Sybil attack nodes in WSN is required to reach 100%. We propose a detection scheme based on residual energy, to further detect the Sybil attack node sneaked through the first-order high-energy nodes. This is occurs on the second-order high-energy nodes.

For each Sybil node, it has several different identities, and it transmits multiple control packets to the nearest first-level high-energy nodes to verify its forged identity. If they get away with the first-level high-energy detection, then first-order high-energy node transmits its control packet to the second-level high-energy node. Suppose ID_i is represent the every forged identities of the one Sybil node, E_i is represent the residual energy of each identity, the whole structure is shown in Fig.2.

In every control packet, it contains the residual energy and the ID of the forged identity. The control packets of sneaked Sybil nodes and other normal nodes are stored in the first-level high-energy, and continues convey to the second-level high-energy nodes. Second-level high-energy receive these control packets begin to compare the E_i, showed in Fig.3.

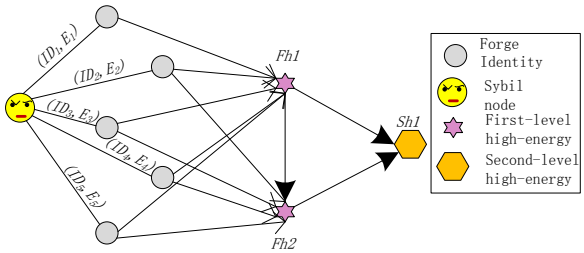

Fig. 2. Sybil nodes and its forged identities.

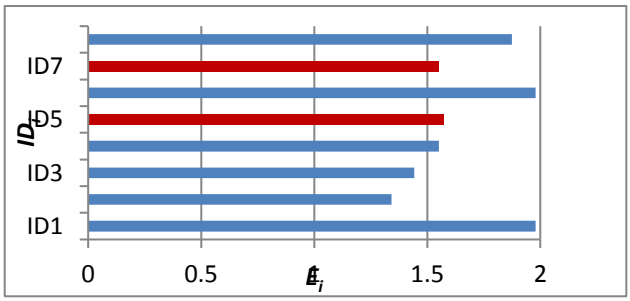

Fig. 3. Residual energy.

From the Fig.3 We can find ID_5 and ID_7 have the similar energy, but they are different identities. So in the each second-level high-energy node, we set a subtracter in each secondary-level high-energy node. In each round, let the residual energy value of ID_1 node be taken as the minuend, and subtract the residual energy value of other nodes respectively. To improve the accuracy, we set the accuracy of the remaining E_i to five decimal places. If the value after subtraction is 0, it indicates that these ID_i are the forged identities from the same Sybil node and Sybil attack occurs. The second-level high-energy nodes will put them to the blacklist and prevent them from compete to the cluster heads.

The second-level detection can greatly improve the detection rate of Sybil node, and make up for the missing detection occurred in the first-level high-energy level. But if our first-level high-energy nodes and second-level high-energy nodes are attacked by Sybil node, the control packets that the second-level high-energy nodes eventually transmit to the base station must have been tampered. To prevent this kind of error operation, another scheme is proposed in the base station.

2.2 Identity match-based scheme

In the base station, relevant information of each normal node and high-level node is stored. Information about Sybil nodes reported by secondary high-energy nodes to the base station is ultimately processed by the base station. Before the base station receives this information, it first checks the first and second high-energy nodes on the link one by one. Further, check whether the identities of these nodes are consistent with those stored on the base station. If not, an attack is considered to have occurred, and the control packet transmitted by this link is unreliable and cannot be received, the base station then blacklists the first and second-level high-energy nodes which attacked by the Sybil nodes. Then base station free the Sybil nodes which stored in those first and second high-energy nodes, let them re-select the link for detection.

Such a three-tier detection mechanism can ensure that every node is detected, and the base station ultimately needs to detect only the first and second-level high-energy nodes, which greatly reduces the pressure of the base station and prolongs the life of the whole network.

3 Experimental results and analysis

In this section, we provide a series of simulation results for our proposed scheme. We use a 200×200 square meter geographical area for our network deployment. Our network comprises of n (n=200) normal nodes and s (from 1 to 30) Sybil nodes, 20 first-level high energy nodes, 10 second-level high energy nodes. Normal nodes and Sybil nodes are isomorphic in terms of battery power, storage and processing capabilities.

We evaluate our scheme in terms of number of detected Sybil nodes, total number of candidates and optimal selection of cluster heads, energy consumption, network lifetime, packet loss rate and packet acceptance ratio.

3.1 Detection of sybil attack

Fig. 4 illustrates the effect of s Sybil nodes and their forged identities on the detection rate for an IWSN. The number of Sybil identities from 5 to 25, n=200. The Sybil nodes and normal nodes have the same residual energy 2J. The first-level high-energy nodes have residual energy 5J, and the second-level high-energy nodes are 7J. In Fig.4, we compared the detection rate of Sybil node with that of [9] and [10]. Due to the three-tier detection of Sybil node in the proposed scheme, its detection rate is relatively high. With the increase of Sybil nodes, the detection rate is more and more ideal.

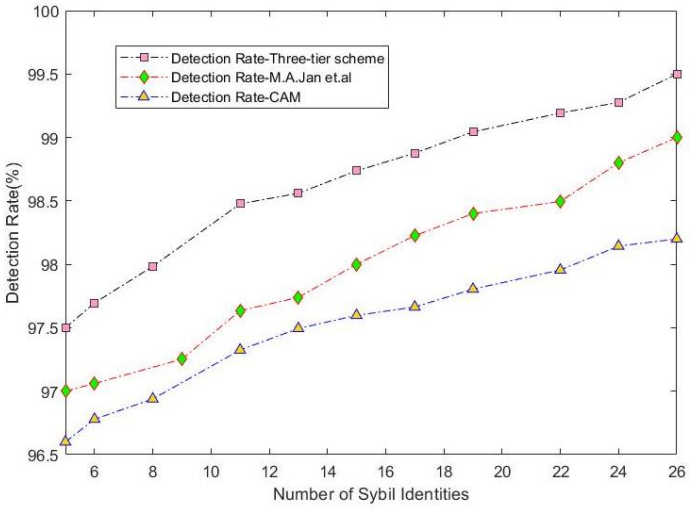

Fig. 4. Residual energy.

3.2 Lifetime of the network

Duo to the limited energy of each node in IWSN, once deployed, it is difficult to replenish energy in the later period. Therefore, the life of the network has become an important issue to consider. Inthe authors believes that when more than 97% of nodes die, the network can no longer carry out follow-up works. Therefore, in Fig.5, we compared the method proposed in

this paper with the method by M.A.Jan and CAM method, mainly to see which scheme can carry more theoretical experiments before the death of 97% nodes. After about 1.7×10^4 rounds, the nodes in our proposed scheme are not available.

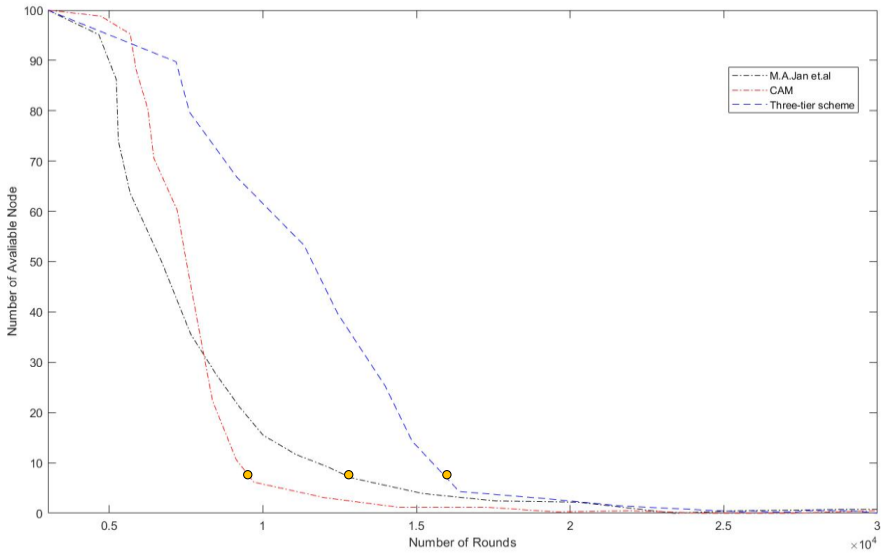

Fig. 5. Lifetime of the network.

3.3 Analysis of energy consumption

The total energy consumption of IWSN is related to the number of Sybil nodes and the number of forged identities. In figure 6, we calculated the energy consumption with and without Sybil nodes. With Sybil nodes, the energy consumption was increased much; the main reason is the Sybil nodes transmitted the control packets. The other reason is the distance from the normal nodes and Sybil nodes to the high energy nodes and the base station. It can lead to the diverse of energy consumption.

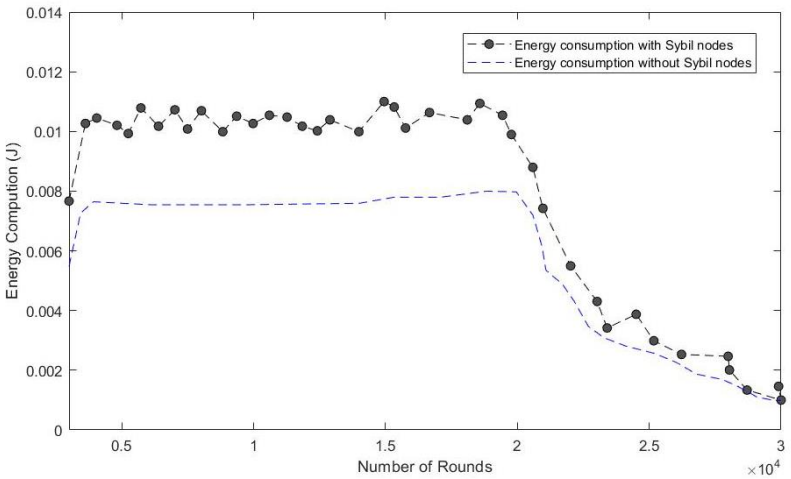

Fig. 6. Energy consumption of the network.

4 Summary

In order to solve the Sybil attack problem in industrial wireless sensor networks, a three-layer detection scheme based on first-level high-energy nodes and second-level high-energy nodes and base station is proposed. Different detection methods are run depending on the size of each level of resources. The first-level high-energy node is detected by the method based on RSSI difference, and the second-level high-energy node is detected by the method based on residual energy. Finally, the base station forms the detection of the first-level high-energy node and the second-level high-energy node. The whole three-tier Sybil attack detection scheme improves the detection of the whole network, improves the detection accuracy, effectively reduces energy consumption and prolongs the network life.

This research is supported by Ordos Institute of applied technology. The project name is “Sybil attack in WSNs forest fire monitoring application”, ID: kyyb2019008.

References

1 J. Yick, B. Mukherjee, D. Ghosal, Wireless sensor network survey, Comput. Netw. 52(2008) 2292-2330.

2 Ashwini I.P, T H Srinivas, A survey on detecting sybil attack in wireless sensor networks, IJARIIE, 3(2), 2017, 163-166.

3 Salam H A, Khan B M. IWSN-standards, challenges and future, IEEE Potentials,35(2), 2016, 9-16.

4 J.R.Douceur, The sybil attack, In First International Workshop on Peer-to-Peer Systems, 02(2002), 1-6.

5 J.Newsome, E,Shi, D.Song, and A.Perring, The Sybil attack in sensor networks: analysis & defenses, Proceedings of the 3rd international symposium on Information processing in sensor networks, ACM, 2004, 259-268.

6 Y.Zhang, W.Liu, W.Lou, and Y.Fang, Location-based compromise-tolerant security mechanisms for wireless sensor networks, IEEE Journal on Selected Areas in Communications, 24(2), 2006, 247-260.

7 K.-F.Ssu,W.-T.Wang, and W.-C. Chang, Detecting sybil attacks in wireless sensor networks using neighboring information, Computer Networks, 53(18),2009,3042-3056.

8 M.Demirbas and Y.Song, An rssi-based scheme for Sybil attack detection in wireless sensor networks, Proceedings of the 2006 Internatinal Symposium on World of Wireless, Mobile and Multimedia Networks, IEEE Computer Society, 2006, 564-570.

9 X.Jin, Research and realization of wireless communication channel propagation model, Beijing university of posts and telecommunications, 2010, 5-18.

10 M.A. Jan, P. Nanda, X. He, R.P. Liu, A sybil attack detection scheme for a centralized clustering-based hierarchical network, in: Trustcom/BigDataSE/ISPA, 2015 IEEE, Vol. 1, IEEE, 2015, 318–325.

11 U.S.Rajkumar, D.Rajamani, A compare and match approach for reventing sybil attacks in wireless sensor networks, IJETSR, 2(2015) 164-172.

12 M.A. Jan, P. Nanda, X. He, R.P. Liu, Pasccc: Priority-based application-specific congestion control clustering protocol, Comput. Netw. 74 (2014) 92–102.