Research on Pseudo-Node Detection Algorithm in Wireless Sensor Networks

https://doi.org/10.3991/ijoe.v13i03.6864

Wenjin Yu
China tobacco Zhejiang Industrial CO., LTD, Ningbo 315040, China
13905746886@139.com

Yong Li
China tobacco Zhejiang Industrial CO., LTD, Ningbo 315040, China
liy@zjtobacco.com

Yuangeng Xu
China tobacco Zhejiang Industrial CO., LTD, Ningbo 315040, China
xuyg163@163.com

Abstract—With the wide application of the wireless sensor network, the security of the sensor network is becoming increasingly important. In this paper, based on node ranging, a new intrusion node detection algorithm has been proposed for external pseudo-node detection in wireless sensor networks. The presence of the nodes under copying-attack and the pseudo-nodes in the network can be detected through inter-node ranging with appropriate use of various sensors of nodes themselves and comprehensive analysis of ranging results. Operating in a stand-alone or embedded manner, this method has remedied the defects in the traditional principle of attack detection. The simulation results show that the proposed method has excellent applicability in wireless sensor security detection.

Keywords—wireless sensor; node algorithm; security detection; application study.

1 Introduction

Wireless Sensor Network (WSN) consists of numerous micro sensor nodes with specific functions in the monitored area, which organize themselves into a wireless network via the wireless communication. It supports the collaborative functions of sensing, collecting and processing the information of the observed objects in network-monitoring area, and transmits the information to the host and the observer [1-3]. These characteristics enable WSN to cover application fields such as military applications, industrial monitoring and control, environmental monitoring, smart home, logistics management, and even anti-terrorism and disaster relief. However, due to the limitations of computing power, communication ability, power supply and storage space of WSN nodes, WSN security mechanism is facing various challenges of secure-communication-oriented design and deployment [4]. Therefore, the research on WSN security is not only of great theoretical significance but also of practical value.

The security target of wireless sensor networks is to ensure the availability, confidentiality, integrity of the network, as well as the authentication and freshness of the nodes [5]. As to node authentication in practical application, the enemy captures sensor nodes for relevant information, and clones the pseudo node into the network in order to obtain, tamper, and forge network information, and even to carry out various malicious attacks on the network and disable the network [6-7]. Aiming at correctly identifying pseudo nodes in sensor networks so as to guarantee the security of WSN, a joint-ranging-based replication attack detection protocol has been designed in this paper to check the existence of pseudo-nodes in WSN by means of inter-node ranging. This method caters for different ranging precision in intrusion detection of replicated nodes in WSN. By adjusting the key contributing parameters to detection rate in combination with the flexible and reliable protocol, the detection success rate can be optimized to meet the requirements of WSN intrusion detection [8-11].

2 Advantages and disadvantages of common node ranging mechanism

Some of the common inter-node ranging mechanisms in WSN include TOA, TDOA, AOA, RSSI. TOA mechanism has good accuracy in inter-node ranging [12]; however, all nodes in the network should be highly synchronized [13]. In addition, higher power consumption of sensor nodes has also been required[14]; Similar to TOA location with little range error and high precision, TDOA mechanism requires a variety of types of sensors, increasing the production cost of sensor nodes and accordingly the energy consumption[15], and thus restricting the application of ranging mechanisms in inter-node ranging in the network; vulnerable to external influences, AOA ranging technology demands extra angular measurement devices which increases the volume and power consumption of sensor nodes. Hence, the mechanism is not available to the inter-node ranging of sensors; based on the received signal strength indication, RSSI ranging technology is not as accurate as the above methods, but it is simple and flexible enough to implement with lower application costs [16-17]. Susceptible to factors like intervisibility condition, temperature humidity, radio environment, communication mode, RSSI ranging mechanism has encountered some bottlenecks in practical system application, albeit well-performed in laboratory test [18].

It is precisely because of the shortcomings of these common node location mechanisms that it is difficult to identify pseudo-nodes in the application of WSN. Therefore, in this paper, a new node-ranging-based intrusion node detection algorithm for WSN has been proposed to better identify and detect pseudo-nodes so as to ensure the security of WSN.

3 Analysis of intrusion node detection algorithm

The system used in the detection algorithm consists of four parts: sensor node, data acquisition terminal, embedded system and data analysis software. Figure 1 shows the composition and architecture of the system.

In wireless sensor networks, as shown in Figure 1, the solid circles represent intrusion nodes, while the normal node is indicated by a hollow circle. Three checking rules have been worked out in accordance with the position relationship between the replay nodes and normal nodes in the network. The data terminal compares and checks the address table of adjacent nodes returned from sensors for any inconsistency between the numbers in the table so as to further determine the presence of intrusion nodes.

Data acquisition terminals and sensor nodes contain multiple functional modules, which are respectively used for wireless communication, data processing, data acquisition, etc., as shown in figure 2.

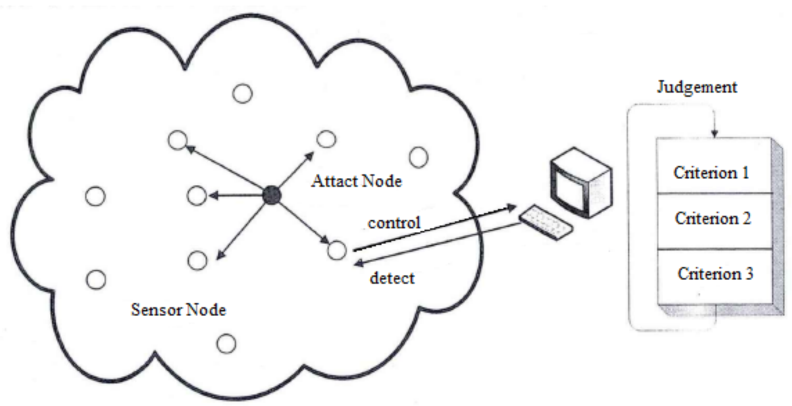

Fig. 1. System architecture diagram

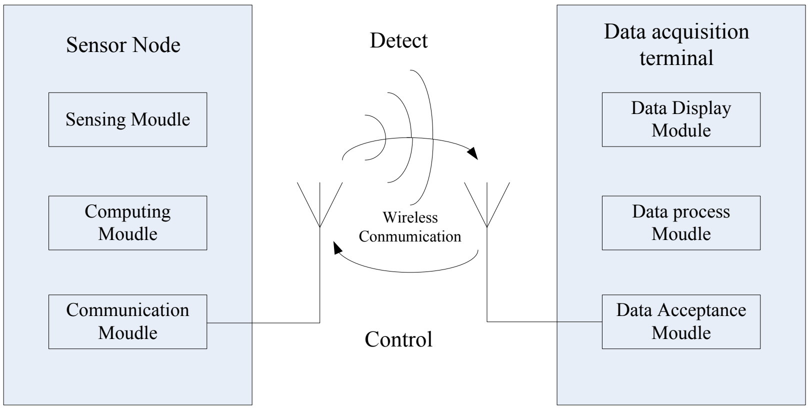

Fig. 2. System modular structure diagram

The sensor module is responsible for indicating signal strength attenuation in communication; the calculation module provides the functionality of processing the collected data and encapsulating them in frames for transmission as well as handling other node tasks and control commands; the communication module takes charge of wireless data transceiver and real-time channel monitoring; the data receiver module receives and processes returned data frames from sensor nodes in order to extract the adjacent node information; the data handling module analyzes the adjacent node information, reads out the ID numbers of intrusion nodes and transmits them to the data display module; finally, the intrusion detection results will be instantly displayed by multi-media through the data display module.

3.1 WSN model hypothesis

Assuming a random WSN application area S, among which is evenly distributed the wireless sensor network with the node size of n. The stable node detection range is *R,* ranging and communication being unaffected by the shape and size of sensor nodes. Each sensor node in the network has a unique node identifier, its number being x_1, x_2…x_n. In the region S, the replicated node *x*' is randomly distributed with the same identity of the node *x*, x_i∈{x_1, x_2…x_n}. In the module, the edge area has a negligible impact on the detection probability of the system, assuming that each node has the same transmission power and electrical characteristics. The node RF chip being equipped with the RSSI data acquisition function or the node being equipped with other inter-node-ranging functional modules, the system provides no accurate node location information and its synchronization time. The inter-node distance measurements are used to detect the attacked-node identification.

Composed of a large number of sensor nodes, WSN achieves inter-node data transmission by means of wireless communication. The joint ranging can be achieved via particular ranging modules. Meanwhile, the RSSI value of the node communication can also be used to calculate the distance between nodes; all the node IDs with a distance of R/2~R and 0~R/2 to each node have been respectively recorded as shown in table 1. By exchanging and comparing the information table of each node, we can determine whether the contradiction exists or whether the node has been invaded.

Table 1. The structure diagram of adjacent-node information table

FLAG	IP
01	6
00	9
00	8
01	6
……	……
00	7

Note:
Number 00 indicate the distance of the nodes in the range of *0~R/2*;
Number 01 indicate the distance of the nodes in the range of *R/2~R*;

3.2 Introduction to three ranging methods

Considering the distributed characteristics of WSN, the protocol only compares the adjacent node information table with less than two nodes. When the node *x* and node *x'* in the system have the same ID, there will be the following three kinds of contradiction:

Self-ranging contradiction: When the replicated node *x'* is within the detection range of node x, i.e. the distance between the two nodes is less than the detection radius R, the contradiction of the same ID will appear in the information table of the two adjacent nodes. In other words, there is a replicated node with the same ID within the detection radius R.

Single node ranging contradiction: When the distance between node *x'* and node *x* is 3R/2~R, there may exist a node that can detect the contradiction between *x'* and *x* in corresponding adjacent node information table; when the distance between node *x'* and node *x* is 3R/2~R, there may exist a node x_k that can detect the contradiction of the same ID number but unequal node range between *x'* and *x* in corresponding adjacent node information table. That is to say, there are two nodes *x* in the area within a distance of *0~R/2* and *R/2~R* from node x_k respectively, and these two nodes are located in their respective areas.

Binode ranging contradiction: When node x_{ki} detects replication node *x'* within the range of 0~R/2 and another node *x* is also in this range, the distance between x_{ki} and $x_{kj}$is larger than *R*, and the replicated node can be detected considering the ranging contradiction.

Being independent of one another, the above three ranging contradictions can be used as independent detective criterions in the protocol. During the detection process of replicated nodes, these three ranging contradictions can be detected in turn. In addition, when finally calculating the detection probability of the protocol, which of the three criterions based on weighting accumulated generating operation should be the total detection probability of the system given the superposition of the detection probability of three criterions.

4 Protocol analysis and Detection Probability Calculation

N nodes are numbered sequentially as x_1, x_2…x_(n-1), set X={x_1, x_2…x_(n-1)}. The distance between node x_i and node x_j is |x_i-x_j |. Assuming that the node captured by the attacker is *x*∈*X* and the replicated node is *x'* ∈*X*, then there is the equation x=x' since the replicated nodes is forged by the captured one. Event W_1, W_2 and W_3 refer to setting the replicated node in accordance with criterion 1, 2 and 3 respectively. Based on the sequential detection method, the probabilities of successful detection are P_1, P_2 and P_3 accordingly. The total probability of being detected by the system for replicated nodes is, and then the equation $P=P_1+P_2+P_3$ is available.

(1) Calculate the probability P_1 of event W_1

∵x_1, x_2…x_(n-1), , maintaining a uniform independent distribution

∴f(x_i) represents the probability that node x_i belongs to a certain point:

$$f(x_i) = \begin{cases} \frac{1}{s}, x_i \in S \\ 0, x_i \notin S \end{cases} \qquad (1)$$

When the replicated node x' is right within the perception range of node x, there is |x-x'|≤R. Therefore, we have:

$$P_1 = \iint_{|x-x'| \le R} f(x) \cdot f(x')\, dS \qquad (2)$$

Plug Formula 1 into Formula 2:

$$P_1 = \iint_{|x-x'| \le R} \frac{1}{s} dS = \frac{\pi R^2}{s} \qquad (3)$$

(2) Calculate the probability P_2 of event W_2 without occurring event W_1

According to the second criterion, there exists a node x_k. Event W_2 represents that x can detect the pseudo node, its probability being P_2. The existence condition of node x_k can be represented as shown:

$$\begin{cases} |x_k \quad x| \le \frac{R}{2}, (A\ Event) \\ \frac{R}{2} < |x_k \quad x'| \le R, (B\ Event) \end{cases} \text{or} \begin{cases} |x_k \quad x'| \le \frac{R}{2}, (A'\ Event) \\ \frac{R}{2} < |x_k \quad x| \le R, (B'Event) \end{cases}$$

The probability P (A) of event A and P (B) of event B can be calculated according to formula 2:

$$P(A) = \iint_{|x_k - x| \le \frac{R}{2}} f(x) \cdot f(x)\, dS = \frac{\pi (R/2)^2}{s} = \frac{\pi R^2}{4s}$$

$$P(B) = \iint_{\frac{R}{2} < |x_k - x'| \le R} f(x_k) \cdot f(x')\, dS = \frac{\pi R^2 - \pi (R/2)^2}{s} = \frac{3\pi R^2}{4s}$$

Event A and event B are independent of each other.

$$\therefore P(AB) = P(A) \cdot P(B) = \frac{\pi R^2}{4s} \cdot \frac{3\pi R^2}{4s} = \frac{3\pi^2 R^2}{16 s^2} \qquad (4)$$

Event AB and event A'B' are mutually symmetric, and P(AB)=P(A'B') accordingly. At this point P_2' can be calculated according to formula 4 without occurring event W_1 :

$$P_2' = [P(AB) + P(A'B')] \cdot (1 - P_1) = \frac{3\pi^2 R^4}{8 s^3} \cdot (s - \pi R^2) \qquad (5)$$

As x_kЄX, the existence of at least one node x_k can be calculated by Bernoulli equation; the probability of satisfying the condition is:

$$P_2 = 1 - (1 - P_2')^{n-1} \tag{6}$$

(3) Calculate the probability *P₃* of event *W₃* without occurring event *W₁* and event *W₂*

Event W_3 indicates the presence of a pair of nodes (x_k1,x_k2)□X that can detect the pseudo node, its probability being P_3. The existence condition of node (x_k1, x_k2) can be represented as shown:

$$\begin{cases} |x_{k1} \quad x| \le \frac{R}{2}.(A\ Event) \\ |x_{k2} \quad x'| \le \frac{R}{2}.(B\ Event) \\ |x_{k1} \quad x_{k2}| > R.(C\ Event) \end{cases} \text{or} \begin{cases} |x_{k1} \quad x'| \le \frac{R}{2}.(A'\ Event) \\ |x_{k2} \quad x| \le \frac{R}{2}.(B'Event) \\ |x_{k1} \quad x_{k2}| > R.(C'Event) \end{cases}$$

Similarly, the probability P (A) of event A, P (B) of event B and P (C) of event C can be calculated according to formula 2:

$$P(A) = \iint_{|x_{k1}-x| \le R/2} f(x) \cdot f(x)\, dS = \frac{\pi(R/2)^2}{s} = \frac{\pi R^2}{4s}$$

$$P(B) = \iint_{\frac{R}{2} < |x_{k2}-x'| \le R/2} f(x_{k2}) \cdot f(x')\, dS = \frac{\pi R^2 - \pi(R/2)^2}{s} = \frac{\pi R^2}{4s}$$

$$P(B) = \iint_{|x_{k1}-x_{k2}| > R} f(x_{k1}) \cdot f(x_{k2})\, dS = 1 - \frac{\pi R^2}{s}$$

Events A, B, and C are mutually independent.

$$\therefore P(ABC) = P(A) \cdot P(B) \cdot P(C) = \frac{\pi R^2}{4s} \cdot \frac{\pi R^2}{4s} \cdot \left(1 - \frac{\pi R^2}{s}\right) = \frac{\pi^2 R^4}{16 s^3}(s - \pi R^2) \tag{7}$$

Event ABC and event A'B'C' are mutually symmetric, and accordingly. At this point P_3' can be calculated according to formula 7 without occurring both event W_1 and W_2:

$$P_3' = [P(ABC) + P(A'B'C')] \cdot (1 - P_1) \cdot (1 - P_2) = \frac{\pi^2 R^4}{8 s^4} \cdot (s - \pi R^2)^2 \cdot (1 - P_2')^{n-1} \tag{8}$$

According to the third criterion, the replicated node can be checked out as long as any set of nodes (x_{k1}, x_{k2}) satisfies the condition. And then based on the Bernoulli equation, the probability P3 can be calculated:

$$P_3 = 1 - (1 - P_3')^{C_n^2 - 1} \tag{9}$$

In view of the above calculation, the copy attack detection probability P based on joint ranging can be obtained:

$$P = P_1 + P_2 + P_3$$

$$= \frac{\pi R^2}{s} + 1 - (1 - P_2')^{n-1} + 1 - (1 - P_3')^{C_n^2 - 1}$$

$$= \frac{\pi R^2}{s} - (1 - P_2')^{n-1} - (1 - P_3')^{C_{n-1}^2} + 2 \tag{10}$$

Where $P_2' = \frac{3\pi^2 R^4}{8s^3} \cdot (s - \pi R^2)$, $P_3' = \frac{\pi^2 R^4}{8s^4} \cdot (s - \pi R^2)^2 \cdot (1 - P_2')^{n-1}$

In the above formula, *P(R, S,n*) represents the application area S, and the functional relationship between R (Detection range) and N (Node scale) has been determined. According to the function relation, the proper selection of S, R and n can ensure the system a higher detection probability of replicated nodes, and then guarantee the reliability and practicability of WSN system.

5 Simulation experiment

This paper adopts NS2.27 as its simulation tool. The experiment runs on a PC processor with a 2.26 GHz dual-core CPU and 4 GB of memory, MAC protocol being IEEE 802.15.4.

To verify the influence of parameters R (Detection range), S (Area) and n (Node number) on the detection rate of the replicated nodes as the system runs, different values of parameters R, S and n were selected in the simulation experiment in order to observe detection probability variation with the three parameters.

(1) Firstly, the simulation results of the influence of parameters R and n on detection probability P are shown in Figure 3 and 4, the detection area S being 10000.

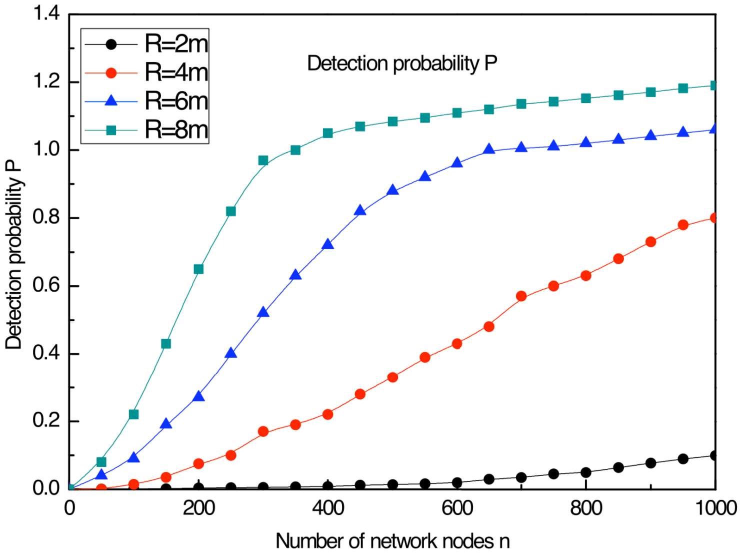

Fig. 3. Curves of detection probability varying with the number of network nodes under different detection distance

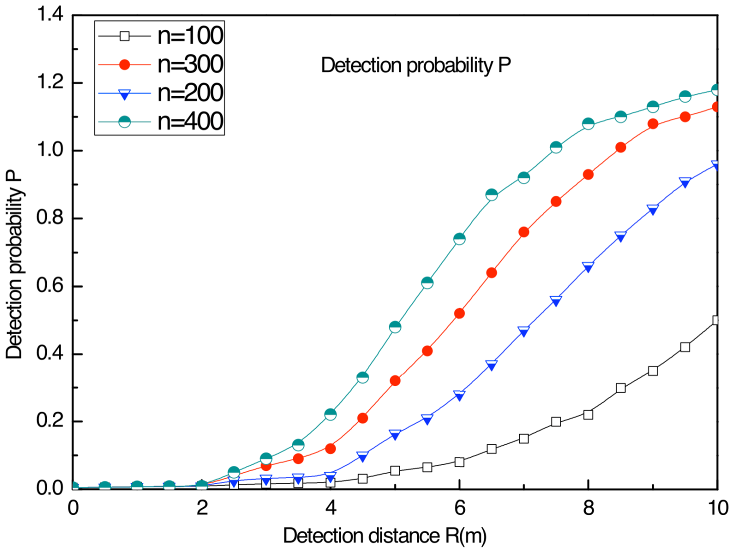

Fig. 4. Curves of detection probability varying with detection distance under different number of network nodes

As shown in figure 4, the detection probability P will reach 100% if system parameters can be selected appropriately according to different application environments. Due to the use of three detection rules which are independent of each other in the detection protocol, the detection probability is greater than 1 within a specific range of detection distance. In this case, the first effective detection rule should be used.

(2) Node detection range R being 3, the detection probability varies with different parameters s and n, as shown in figure 5.

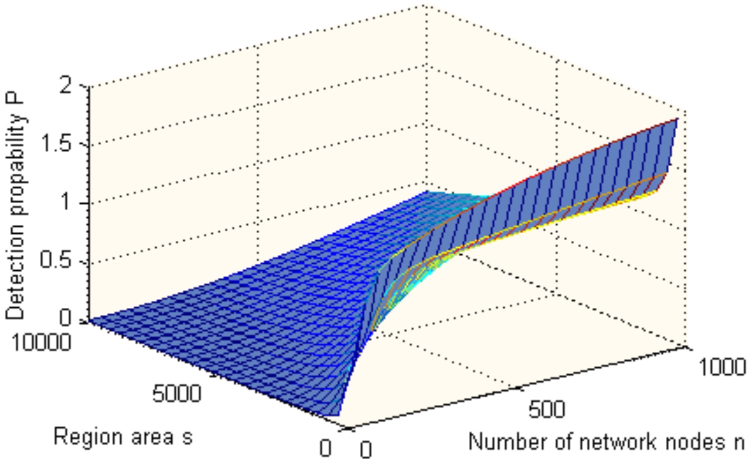

Fig. 5. Detection probability variation with node detection range R being 3

As shown in figure 5, the system detection probability approaches and reaches up to 100% along with the change of parameters S and n. The darker part of the surface indicates smaller detection probability; the lighter part indicates larger detection probability. Therefore, when the node detection range is fixed in the system, the detection probability demand can be achieved by configuring and verifying the system node size n according to the detection area S.

(3) Next, set node coverage area S to be 10000, and the simulation results of detection probability variation with the change of parameters R and n are shown below in Figure 6:

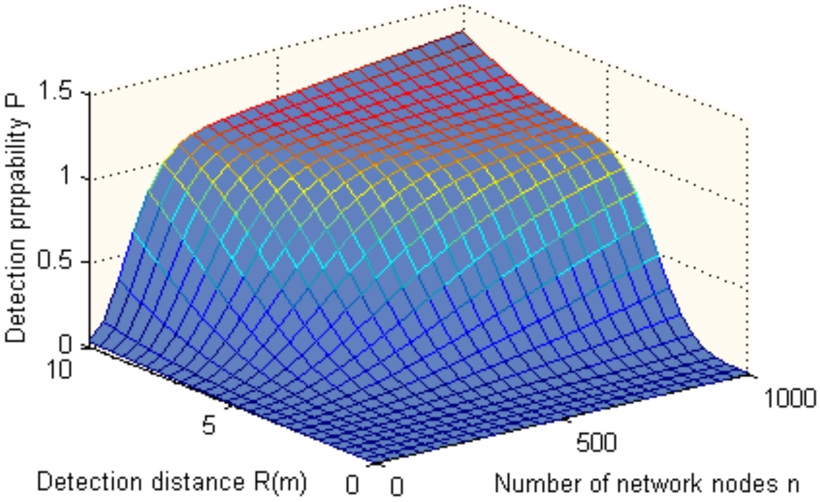

Fig. 6. Detection probability variation with coverage area S being 10000

The experimental results show that, with fixed coverage area, the radius of node R and node size n have a significant impact on the detection probability P. The appropriate selection of parameters R and n ensures the detection probability of the system.

It can be seen from the above simulation experiment that all the protocols adopted for R, S and n of varying values can achieve higher pseudo node detection probability. Moreover, the node-ranging-based intrusion node detection algorithm for WSN, being applicable for barely functional wireless sensor nodes, greatly contributes to the cost control of WSN system. At the same time, high replay node detection probability can be achieved if the application area is equipped with sensors and nodes with appropriate scale. In addition, the algorithm can obtain the intrusion node ID by measuring the signal intensity attenuation without the synchronizing information of node position and system clock.

During its operation, the algorithm can identify the intrusion node by means of self-adaptive methods and those three judging criterions. The strength value of radio frequency signal is used as a decision in the system, which not only ensures a high recognition rate of pseudo nodes but also makes the algorithm maintain high adaptability. Therefore, the algorithm can be widely adopted in WSN and other networks.

6 Conclusion

This paper introduces the system structure of the node-ranging-based intrusion node detection algorithm for WSN; it also describes the method and basis of intrusion node detection through the analysis of ranging methods and joint-ranging-based copy attack detection protocol. Besides, the feasibility of the algorithm is verified by simulation experiments. The protocol principle of this paper is simple and reliable, without additional hardware, high cost or high power consumption. The methods based on node ranging can effectively detect replicated nodes; meanwhile, the system parameters in the protocol not only provide basis for sensor node selection and system scale setting but also optimize the detection probability, reducing the presence probability of pseudo nodes in WSN and enhancing the security of WSN.

7 References

[1] Luo, J., Hu, J., Wu, D., Li, R. (2015). Opportunistic routing algorithm for relay node selection in wireless sensor networks. IEEE Transactions on Industrial Informatics, 11(1), 112-121. https://doi.org/10.1109/TII.2014.2374071

[2] Huang, X., Zhang, X. (2013). A node deployment algorithm based on van der waals force in wireless sensor networks. International Journal of Distributed Sensor Networks, 2013(2013), 485-503.

[3] Chen, X., Zhang, B. (2012). Improved dv-hop node localization algorithm in wireless sensor networks. International Journal of Distributed Sensor Networks, 2012(6), 1018-1020.

[4] Muruganathan, S. D., Ma, D. C. F., Bhasin, R. I., Fapojuwo, A. O. (2005). A centralized energy-efficient routing protocol for wireless sensor networks. IEEE Communications Magazine, 43(3), S8-13. https://doi.org/10.1109/MCOM.2005.1404592

[5] Yan, J., Ling, W., Yang, X. Z., Wen, D. X. (2007). Overview of node scheduling algorithm in wireless sensor networks. Yuhang Xuebao/journal of Astronautics, 28(5), 1086-1093.

[6] Ma, C., Liang, W., Zheng, M., Sharif, H. (2015). A connectivity-aware approximation algorithm for relay node placement in wireless sensor networks. IEEE Sensors Journal, 16(2), 515-528. https://doi.org/10.1109/JSEN.2015.2456931

[7] Chen, Y., Lu, S., Chen, J., Ren, T. (2016). Node localization algorithm of wireless sensor networks with mobile beacon node. Peer-to-Peer Networking and Applications, 1-13.

[8] Xiao, W., Wu, X., Ma, X., Lu, Q. (2005). Optimization algorithm of wireless sensor network node, improved ant colony, ant colony optimization algorithm, network node, wireless sensor network, tiny sensor nodes, wireless sensor networks, wireless sensor network node. International Journal of Distributed Sensor Networks, 20(4), 1-4.

[9] Zda, R., Karc, A. (2015). Sensor node deployment based on electromagnetism-like algorithm in mobile wireless sensor networks. International Journal of Distributed Sensor Networks, 2015, 1-15.

[10] Srivastava, J. R., Sudarshan, T. S. (2015). Energy-efficient cache node placement using genetic algorithm in wireless sensor networks. Soft Computing, 19(11), 3145-3158. https://doi.org/10.1007/s00500-014-1473-8

[11] Fan, G. J., Wang, R. C., Huang, H. P., Sun, L. J. (2011). Tolerable coverage area based node scheduling algorithm in wireless sensor networks. Tien Tzu Hsueh Pao/acta Electronica Sinica, 39(1), 89-94.

[12] Nokhanji, N., Hanapi, Z. M., Subramaniam, S., Mohamed, M. A. (2015). An energy aware distributed clustering algorithm using fuzzy logic for wireless sensor networks with non-uniform node distribution. Wireless Personal Communications, 84(1), 1-25. https://doi.org/10.1007/s11277-015-2614-9

[13] Zhang, Y., Xiang, S., Fu, W., Wei, D. (2014). Improved normalized collinearity dv-hop algorithm for node localization in wireless sensor network. International Journal of Distributed Sensor Networks, 2014(11), 1-14. https://doi.org/10.1504/IJSNET.2014.065778

[14] Xiao, F., Wu, M., Huang, H., Wang, R., Wang, S. (2012). Novel node localization algorithm based on nonlinear weighting least square for wireless sensor networks. International Journal of Distributed Sensor Networks, 2012(2012), 1238-1241.

[15] Jing, H. (2015). Node deployment algorithm based on perception model of wireless sensor network (special issue on production planning and scheduling). International Journal of Automation Technology, 9, 210-215. https://doi.org/10.20965/ijat.2015.p0210

[16] Fouad, M. M., Snasel, V., Hassanien, A. E. (2015). Energy-aware sink node localization algorithm for wireless sensor networks. International Journal of Distributed Sensor Networks, 2015(1), 134. https://doi.org/10.1155/2015/810356

[17] Chen, X., Chen, J., Chen, C., He, J., & Lei, B. (2013). A node localization algorithm for wireless sensor networks using distance clustering to select the anchor nodes. Sensor Letters, 11(4), 745-748. https://doi.org/10.1166/sl.2013.2512

[18] Wei, G., Yan, S., Fan, S. (2015). Routing algorithm based on area division management of node in wireless sensor networks. Telkomnika, 13(4), 1214. https://doi.org/10.12928/telkomnika.v13i4.1895

8 Authors

Wenjin Yu, male, master, senior engineer. He engaged in manufacturer information research and management for a long time, his research mainly focused on researching enterprise information management, production network and production technologies. He is with China tobacco Zhejiang Industrial CO., LTD, Ningbo 315040, China (13905746886@139.com).

Yong Li, male, bachelor, senior engineer. He engaged in enterprise information construction and project management. His research mainly focused on enterprise information management and process automation. He is with China tobacco Zhejiang Industrial CO., LTD, Ningbo 315040, China (liy@zjtobacco.com).

Yuangeng Xu, male, bachelor, engineer. He engaged in production process automation, network construction and project management. His research mainly focused on digital factory and process automation. He is with China tobacco Zhejiang Industrial CO., LTD, Ningbo 315040, China (xuyg163@163.com).

Article submitted 20 February 2017. Published as resubmitted by the authors 23 March 2017.