Journal of Zhejiang University SCIENCE A
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

WAPN: a distributed wormhole attack detection approach for wireless sensor networks

Fan-rui KONG[†1], Chun-wen LI[1,3], Qing-qing DING[2], Guang-zhao CUI[3], Bing-yi CUI[4]

([1]*Department of Automation, Tsinghua University, Beijing 100084, China*)

([2]*State Key Lab of Power Systems, Department of Electrical Engineering, Tsinghua University, Beijing 100084, China*)

([3]*State Key Lab of Informational Electric Apparatus in Henan, Zhengzhou 450002, China*)

([4]*Department of Physics, Nanjing University, Nanjing 210093, China*)

[†]E-mail: kongfr@mails.tsinghua.edu.cn

Received Mar. 13, 2008; Revision accepted June 6, 2008; Crosschecked Dec. 26, 2008

Abstract: As the applications of wireless sensor networks (WSNs) diversify, providing secure communication is emerging as a critical requirement. In this paper, we investigate the detection of wormhole attack, a serious security issue for WSNs. Wormhole attack is difficult to detect and prevent, as it can work without compromising sensor nodes or breaching the encryption key. We present a wormhole attack detection approach based on the probability distribution of the neighboring-node-number, WAPN, which helps the sensor nodes to judge distributively whether a wormhole attack is taking place and whether they are in the influencing area of the attack. WAPN can be easily implemented in resource-constrained WSNs without any additional requirements, such as node localization, tight synchronization, or directional antennas. WAPN uses the neighboring-node-number as the judging criterion, since a wormhole usually results in a significant increase of the neighboring-node-number due to the extra attacking link. Firstly, we model the distribution of the neighboring-node-number in the form of a Bernoulli distribution. Then the model is simplified to meet the sensor nodes' constraints in computing and memory capacity. Finally, we propose a simple method to obtain the threshold number, which is used to detect the existence of a wormhole. Simulation results show that WAPN is effective under the conditions of different network topologies and wormhole parameters.

Key words: Wireless sensor networks (WSNs), Security, Wormhole detection
doi:10.1631/jzus.A0820178 **Document code:** A **CLC number:** TP393

INTRODUCTION

A wireless sensor network (WSN) is a self-organized network, consisting of hundreds or thousands of sensor nodes (we use the terms 'sensor node', 'sensor' and 'node' interchangeably in this paper) (Sohrabi *et al.*, 2000; Qi *et al.*, 2001; Aboelaze and Aloul, 2005). The development of low-power and low-cost electrical devices makes WSNs well suited for use in various military and civil applications (Akyildiz *et al.*, 2002; Chong and Kumar, 2003; Faza and Sedigh, 2006). In most military applications, WSNs are used in hostile areas, where nodes can be compromised and the channel can be deliberately jammed by the enemy. In some civil applications, personal and private data need to be protected from leaking. So communication security is a crucial issue and deserves special attention during the design process (Slijepcevic *et al.*, 2002; Boudriga and Obaidat, 2006).

There are many types of attacks such as packet injection, miscellaneous attacks against routing, Sybil attack, wormhole attack, and hello flood attack (Shi and Perrig, 2004; Roman *et al.*, 2005). Wormhole attack is a special kind of outside attack, which can result in severe damage to the functions and structures of WSNs, especially to the routing scheme (Karlof and Wagner, 2003). In a wormhole attack, the attacker places two special nodes in the sensing area, a malicious node and a colluding node. The malicious

node overhears and tunnels all the data it receives to the colluding node via a low-latency, high-quality, and out-of-bound channel (the attacking tunnel) available only to the attackers. The colluding node then replays the data in the channel used by all the sensor nodes at its position. The attacking tunnel is usually asymmetric. The colluding node forwards packets selectively back to the malicious node or forwards no packet at all. Therefore, a wormhole attack can lead to invalidation of the routing protocol by misdirecting the data flow, selective forwarding, or disconnecting a certain area from the whole network (Karlof and Wagner, 2003). The fatal feature of a wormhole attack is that it can attack the network without compromising sensor nodes or breaching the cryptography of the network.

In this paper, we present a distributed wormhole attack detection approach, WAPN, based on the observation that a sensor node affected by a wormhole usually suffers a sharp increase in its neighboring-node-number due to the attacking tunnel, which can lead two nodes located more than one hop away into believing that they are immediate neighbors. WAPN has two important features:

(1) It differs from previous detection approaches in that WAPN can be easily implemented in resource-constrained WSNs without any additional requirements, such as node localization, tight synchronization, or directional antennas.

(2) WAPN is scalable. WAPN becomes more accurate as the scale of the network increases.

RELATED WORK

Poovendran and Lazos (2007) presented a graphic theoretical approach for modeling wormhole links. They showed that any possible countermeasure against the wormhole attack should construct a communication graph that is a connected subgraph of the geometric graph of the network.

In (Hu *et al*., 2003), two kinds of packet mechanisms, geographical leashes and temporal leashes, were used to detect the wormhole attack. In the geographical leashes mechanism, each node must know its own location, while in the temporal leashes mechanism the whole network needs to be tightly synchronized. However, neither the localization of the sensor nodes nor network synchronization can be easily achieved in WSNs.

Buttyan *et al*.(2005) used the neighbor number test (NNT) to detect an increase in the number of the neighbors and the all distances test (ADT) to detect a decrease in the lengths of the shortest paths between all pairs of sensors. NNT is based on the distribution pattern of the neighboring-node-number, which is similar to our model. But the detection algorithm needs the sink node to do a χ^2 test and to keep two histograms of neighboring information on all nodes. When the node number is large, the algorithm will require the sink to have a high computing and memory capacity. Buttyan *et al*.(2005) assumed that the sinks have no resource limitations and can run complex algorithms. But this cannot be satisfied in some applications of WSNs, where the sink is also resource constrained.

In (Hu and Evans, 2004), a detection scheme based on directional antennas was proposed. Sensor nodes with directional antennas can tell from which direction the packets come, so the direction in which the sender sends the packet must be opposite to the direction in which the receiver receives the packet. If a wormhole attack takes place, bogus neighbors (introduced by the wormhole) cannot always maintain such features except when the nodes happen to reside in opposite directions. Hu and Evans (2004) compensated for this with the help of a verifier, a neighboring sensor node in a direction other than those of the pair of nodes sending and receiving. The approach presented in (Hu and Evans, 2004) is easy to compute and can be engaged distributively, but it needs directional antennas, which are not always available in WSNs. Furthermore, the overhead introduced by communications between nodes and their verifiers may lead to energy squander.

In (Wang and Bhargava, 2004), a wormhole detection mechanism, MDS-VOW, was proposed. Sensor nodes estimate the distance between themselves and their neighboring nodes, and then send the distance information to the sink. The sink reconstructs the layout of the sensors based on this distance information. When a wormhole attack takes place, the reconstructed surface will bend due to the fake connections through the wormhole. By detecting the bending feature, the wormhole is located and the fake connections are identified. Wang and Bhargava

(2004)'s work is another centralized scheme like Buttyan *et al.*(2005)'s work. The algorithm requires high computing and memory capacities in the sink.

Qian *et al.*(2007) proposed a wormhole detection approach, SAM, for WSNs operating under multi-path routing. The attacking tunnel is always more attractive in most multi-path routing protocols like SMR (Lee and Gerla, 2001) due to its low latency and good quality. So certain routes will be adopted more frequently than others. SAM is based mainly on the observation that certain statistics of the routes discovered by routing protocols will change dramatically under wormhole attacks. Hence, it is possible to examine such statistics to detect this type of wormhole attack and pinpoint the attackers if enough routing information is available. Qian *et al.*(2007)'s work is quite effective and accurate in detecting the wormhole attack with the help of multi-path routing, but cannot work under uni-path routing protocols.

Compared with (Hu *et al*., 2003; Hu and Evans, 2004), our approach does not need extra node or network function such as localization of the sensor nodes or synchronization in the network. Unlike (Wang and Bhargava, 2004; Buttyan *et al*., 2005), WAPN is a distributed approach. Our approach is similar to that in (Qian *et al.*, 2007), but the detection criteria are different. We use the statistic feature of the neighboring-node-number to detect the wormhole attack, since when a wormhole attack takes place the neighboring-node-number shows a sharp increase in the area near the colluding node.

ASSUMPTIONS, NOTATIONS AND MODELS

Network model

In a typical application of WSNs, sensors obtain data from the environment and send it to the sink, where the data are integrated, compressed, and transmitted to the users by infrastructures like LAN (local area network), the Internet, or satellites. The network model used in our detection approach is depicted in Fig.1. It employs the following assumptions:

(1) All the nodes are static, including the sensor nodes and the sinks.

(2) Sensor nodes are homogeneous with a fixed radio range *R*.

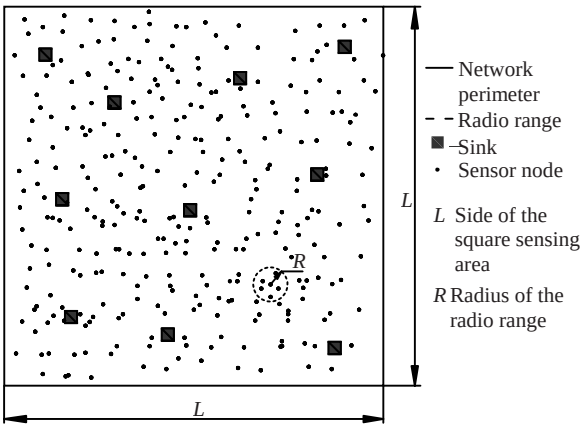

Fig.1 The network model used in our detection approach

(3) All the nodes are randomly and uniformly disposed in a square sensing area whose side is *L*.

(4) The sensing area is relatively broad compared with the radio range of a single sensor node, which means $L>>R$.

(5) Sensor nodes are redundantly disposed. The total number of the nodes in the network is *n*.

(6) The sensor nodes can tell who are their neighbors and what is the number of their neighboring nodes.

These assumptions are widely applied in many research studies and are also practical in real WSN applications. Sensor nodes are mostly low-cost and constrained in energy, computing and communication resources. Due to their energy constraints, sensor nodes are prone to invalidation when the energy supply becomes depleted. Redundant disposal is often used to compensate for the invalidation of the sensor nodes and to prolong the network lifespan. Sensor nodes can easily tell who their neighbors are, since the recognition of the neighboring nodes can be realized by backing the packets with the information of the sending node, or by broadcasting a hello packet periodically.

Wormhole attack model and security assumptions

A wormhole is a kind of outside attack in which the adversary can accomplish the attack without breaching the encryption key or needing the help of a compromised node. In a wormhole attack, the adversary sets two nodes in the network and builds up a low-latency and high-quality link between them. The link is referred to as the 'attacking tunnel' and is

available to only these two nodes. One node, referred to as the 'malicious node', overhears the packets and sends them via the attacking tunnel to the other node, referred to as the 'colluding node'. The colluding node replays the packets at its position. So the nodes in the radio range of the colluding node can hear the nodes whose radio range covers the malicious node, and take these nodes as their neighbors. Our wormhole attack model is depicted in Fig.2. The distance between the malicious node and the colluding node is *D*. The colluding node replays all the packets from the malicious node in the sensor nodes' channel with the radio range R_w. It is assumed that the colluding node does not forward any packets from the sensor nodes back to the malicious node. We define the area one hop away from the colluding node as the 'influencing area' of the wormhole attack. The sensor nodes in the influencing area are referred to as 'poisoned nodes'.

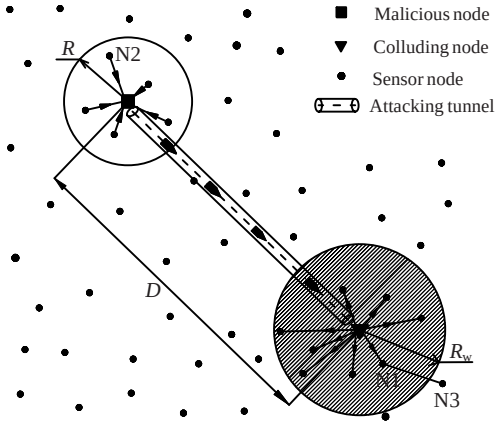

Fig.2 Our wormhole attack model

The routing path using the attacking tunnel seems to be more attractive to the poisoned nodes because it has low latency and few hops, and is of high quality. For example, a poisoned node, N1, can hear the sensor node whose radio range covers the malicious node, N2, and may take N2 as its next hop node due to the advantage of the attacking tunnel (Fig.2). However, N1 and N2 are not real neighbors. The packets destined for N2 from N1 cannot reach N2 because the attacking tunnel is unidirectional. As a result of the wormhole attack, not only the poisoned node may suffer such traffic jamming, but also the nodes near the influencing area. N3, which takes N1 as its next hop node, will suffer the same traffic jamming as N1. Thus there will be a routing confusion area, where the sensor nodes may misaddress their packets and cannot send their packets to the sink. Such a routing confusion area may extend to several hops beyond the colluding node. However, if the poisoned node can self-identify the wormhole attack and broadcast a warning packet to inform its neighbors not to take it as their next hop node, the extension of the routing confusion area can be effectively limited, and then the damage caused by the wormhole attack can be minimized.

Cryptography and message authentication code are widely used as security primitives for WSNs. Public-key cryptography, which requires more resource in computing and memory but provides higher-level security, is also under research (Gura *et al.*, 2004; Doyle *et al.*, 2006). Due to resource constraints, the complexity of encryption algorithms used in WSNs must be appropriate. This research is important but beyond the scope of this paper. It is assumed in WAPN that certain encryption algorithms and key infrastructures are used, so a single node can securely get the conversation key with any other node and the authenticated key to broadcast in its radio range. The encryption algorithm is also assumed to be strong enough. The malicious node and the colluding node cannot modify the packets, so the correctness and integrity of the data on the packets can be assured.

Stochastic model for neighboring-node-number distribution

As in the network model introduced and depicted in Fig.1, the sensor nodes are uniformly disposed in the sensing area. The coordinates *x* and *y* of each sensor node both conform to uniform distributions, which are also independent. So the probability density functions of *x*, *y* and the joint distribution of (*x*, *y*) are

$$\begin{cases} \psi(x)=\begin{cases}1/L, & 0\le x\le L,\\ 0, & \text{else},\end{cases} \\ \psi(y)=\begin{cases}1/L, & 0\le y\le L,\\ 0, & \text{else},\end{cases} \\ \psi(x,y)=\begin{cases}1/L^2, & 0\le x\le L,\ 0\le y\le L,\\ 0, & \text{else}.\end{cases}\end{cases} \tag{1}$$

For a given sensor node A, whose coordinate is (x_A, y_A), the probability of an arbitrary node B being A's neighbor, p_A, varies with (x_A, y_A). It is obvious that p_A is smaller when node A is in the peripheral part of the square sensing area than in the center part, as shown in Fig.3, since the radio range of the node in the peripheral part extends beyond the sensing area. Under the assumption of $L>>R$, such a difference can be ignored, because the nodes in the peripheral part occupy only a tiny amount of the overall nodes.

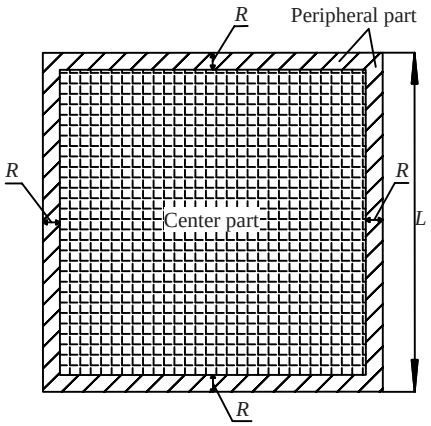

Fig.3 Center part and peripheral part of the sensing area

If we ignore the difference resulting from the position of the sensor node, for sensor node A,

$$p_A = \iint_{G(A)} \psi(x,y)\mathrm{d}x\mathrm{d}y = \frac{1}{L^2}\iint_{G(A)} \mathrm{d}x\mathrm{d}y = \frac{\pi R^2}{L^2}, \quad (2)$$

where G(A) is a circle with a radius of R and node A at its center. Eq.(2) shows that p_A is independent of A's coordinates, which means that the probability of any two nodes being neighbors is $p=\pi R^2/L^2$.

Since all sensor nodes are disposed uniformly and independently in the sensing area, the probabilities of any two nodes being neighbors are also independent. Therefore, the neighboring-node-number of an arbitrary sensor node conforms to a Bernoulli distribution. The probability of a node having i neighbors, $f(i)$, is

$$f(i) = \binom{n}{i} p^i (1-p)^{n-i}, \quad (3)$$

where n is the number of nodes.

WORMHOLE DETECTION APPROACH

The sensor nodes in the influencing area, the poisoned nodes, would suffer an abnormal increase in the neighboring-node-number because of the extra link (the attacking tunnel) introduced by the wormhole. For example, in Fig.4, the malicious node is in the radio range of sensor nodes M1, M2, M3, M4 and M5. C8 is a node in the influencing area and has five real neighbors, C6, C7, N1, N2 and N3. Due to the wormhole attack, C8 can also receive the packets originating from M1, M2, M3, M4 and M5 via the attacking tunnel. Thus, C8 has another five bogus neighbors. The main idea of WAPN is to help the sensor node to self-identify whether it is in the influencing area, based on analyzing and computing the probability of the abnormal increase in the neighboring-node-number. WAPN uses a threshold neighboring-node-number T to detect the existence of the wormhole attack and whether the sensor node is in the influencing area. A node asserts that a wormhole attack exists and the node is in the influencing area of the wormhole, only if the node has more than T neighboring nodes.

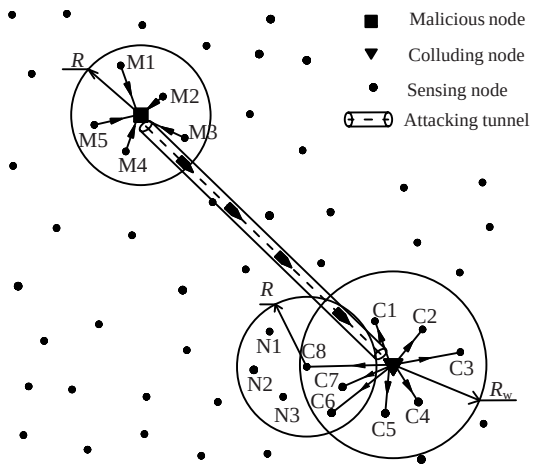

Fig.4 Poisoned nodes' abnormal increase in their neighboring-node-number

We define the possibility of a certain node having less than k neighboring nodes as $F(k)$,

$$F(k) = \sum_{i=0}^{k-1} f(i) = 1 - \sum_{i=k}^{n} f(i). \quad (4)$$

$f(i)$ is defined in Eq.(3). If the value of $F(k)$ approximates to 1, we set k as the neighboring node threshold number T, since there is little possibility that a normal node has more than or equal to k neighboring nodes. We define a critical possibility α to quantify the approximation of $F(k)$ to 1. So we can use α to compute T, which satisfies

$$F(T) > \alpha. \tag{5}$$

It is hard for the resource-constrained sensor nodes to obtain T by directly solving Eq.(5), because the finite series $F(k)$ does not have a closed-form solution. As a result, we first simplify the expression of $F(k)$ and then try to find a lower bound of it, referred to as $\Phi(k)$. We set T as the smallest integer k that satisfies $\Phi(k)>\alpha$. Since $\Phi(k)$ is a lower bound of $F(k)$, T can also satisfy $F(T)>\alpha$. $\Phi(k)$ is a closed-form expression, and solving the inequality $\Phi(k)>\alpha$ to obtain the value of T is simple enough to implement on the sensor nodes.

Note that, to achieve both a high identification rate of the poisoned nodes and a low misjudgment rate of the normal nodes, the value of α should be appropriate. The impact of α on the performance of WAPN will be analyzed in detail in the following.

Analysis and simplification of $F(k)$

From Eq.(3), we have

$$f(k+1) - f(k) = \frac{f(k)}{(k+1)(1-p)}(np + p - 1 - k). \tag{6}$$

So

$$f(k+1) - f(k) = 0 \Rightarrow k = np + p - 1. \tag{7}$$

We define $M = \lceil np + p - 1 \rceil$, which is the smallest integer greater or equal to $np+p-1$, so $M - 1 < np + p - 1 \le M$, and

$$f(k+1) - f(k) \begin{cases} > 0, & k \le M - 1, \\ \le 0, & k \ge M. \end{cases} \tag{8}$$

So the Bernoulli distribution like Eq.(3) is maximized at M. Define

$$L(k) = \frac{f(k+1)}{f(k)} = \frac{(n-k)p}{(k+1)(1-p)}, \tag{9}$$

and then for arbitrary $i \ge k+1$,

$$f(i) = f(k)\prod_{j=k}^{i-1} L(j). \tag{10}$$

It is obvious that $L(k)$ is strictly decreasing as k increases, so for arbitrary $j \ge k$,

$$L(j) \le L(k). \tag{11}$$

Substituting Eq.(11) into Eq.(10), we have for arbitrary $i \ge k+1$

$$f(i) = f(k)\prod_{j=k}^{i-1} L(j) \le f(k)(L(k))^{i-k}. \tag{12}$$

Therefore

$$\sum_{i=k}^{n} f(i) \le f(k)\sum_{i=k}^{n}(L(k))^{i-k} \le f(k)\sum_{i=k}^{\infty}(L(k))^{i-k}. \tag{13}$$

Since $f(k)$ is maximized at M and $L(k)$ is strictly decreasing, we can obtain

$$L(M) = \frac{f(M+1)}{f(M)} \le 1,\ L(k) < L(M) \le 1,\ \forall k > M. \tag{14}$$

So for $k>M$, Eq.(13) can be further simplified as

$$\sum_{i=k}^{n} f(i) \le f(k)\sum_{i=k}^{\infty}(L(k))^{i-k} = f(k)\frac{(k+1)(1-p)}{k+1-p-np}. \tag{15}$$

Finally

$$F(k) = 1 - \sum_{i=k}^{n} f(i) \ge 1 - f(k)\frac{(k+1)(1-p)}{k+1-p-np}. \tag{16}$$

Computation of the threshold T

As shown in Eq.(15), the upper bound of $\sum_{i=k}^{n} f(i)$ is relevant to $f(k)$. Sensor nodes are redundantly disposed in most WSNs, so n is usually several hundred or thousand. According to the assumption $R \ll L$ and $p \ll 1$, the Bernoulli distribution approximates a Poisson distribution:

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k} \approx \frac{(np)^k}{k!} e^{-np}. \tag{17}$$

For $p<<1$,

$$\frac{(k+1)(1-p)}{k+1-p-np} \approx \frac{k+1}{k+1-np}. \quad (18)$$

Therefore

$$f(k)\frac{(k+1)(1-p)}{k+1-p-np} \approx \frac{(np)^k}{k!}\mathrm{e}^{-np}\frac{k+1}{k+1-np}. \quad (19)$$

We define $\Phi(k)$ as

$$\Phi(k) = 1-\frac{(np)^k}{k!}\mathrm{e}^{-np}\frac{k+1}{k+1-np},\ k > M. \quad (20)$$

Substituting Eqs.(19) and (20) into Eq.(16), we can obtain

$$F(k) \geq \Phi(k),\ k > M. \quad (21)$$

np is the mathematical expectation of a Bernoulli distribution like Eq.(3), and is also the average number of the sensor nodes within a circle of radius R in the sensing area. To maintain a certain extent of redundancy, np is usually set as a constant before setting up the network and is a prerequisite for the determination of the number of the nodes disposed in the network. So given the value of α, the sensors can solve the inequality

$$\Phi(k) > \alpha \quad (22)$$

to obtain T with the assumption that np is a constant.

Selection of the critical possibility α

The critical possibility, α, has a great influence on the performance of WAPN. As shown in Eq.(22), T is strongly related to α and α should be given before the network and the sensors are deployed [see Appendix for the relationships between T and α, and between $F(T)$ and α with different values of np]. From Tables A1 and A2 in Appendix, we find that given α=0.8, $F(T)$ is more than 0.99, which is large enough to guarantee a convincible identification rate. In the following section, we set α=0.8 for all the simulation cases.

SIMULATION AND PERFORMANCE EVALUATION

Simulation setup

We used the identification rate r_{ide} and the error rate r_{err} to evaluate the performance of WAPN. r_{ide} is the ratio of the number of poisoned nodes that can accurately identify themselves as poisoned nodes to the overall number of poisoned nodes, and it represents the detection accuracy of WAPN. r_{err} is the ratio of the number of normal nodes that identify themselves as poisoned nodes to the overall number of sensor nodes in the network, and it represents the extent of the hazard caused by mis-identification using WAPN. We define by S_1 the set of the nodes that are actually poisoned, S_2 the set of the nodes that identify themselves as poisoned ones (the nodes that have more than T neighbors), $S_3=S_1\cap S_2$ the set of the nodes that accurately identify themselves as poisoned ones, and n_1, n_2, n_3 the element number of the set S_1, S_2, S_3, respectively. So we have

$$r_{\mathrm{ide}} = n_3 / n_1,\quad r_{\mathrm{err}} = (n_2 - n_3)/n.$$

In order to evaluate the performance of WAPN under different network topologies and wormhole features, we built a simulator with four variable parameters: R—radio range of the nodes, n—number of the nodes in the network, R_{w}—radio range of the colluding node, D—distance between the malicious node and the colluding node, and two constant parameters: L—side of the sensing area, α—critical possibility.

The values of these parameters are listed in Table 1. Each simulation case comprised a particular combination of the parameters. For each simulation case we completed 100 runs and averaged the results of r_{ide} and r_{err} to evaluate the performance of WAPN.

Performance of WAPN under different network topology parameters, n, and R

Node redundancy is a critical feature of WSNs. An important performance evaluation criterion of WSNs is whether an approach or scheme can still maintain its advantage as redundancy increases. Therefore, we ran the simulation with various values of n under the condition that D=500 m, R_{w}=60 m, α=0.8. The results are shown in Fig.5.

Table 1 Parameters of the simulation

Parameter	Possible value
R (m)	40, 50, 60, 70, 80
R_w (m)	20, 40, 60, 80, 100
D (m)	100, 300, 500, 700, 900, 1000, 1300
n	500, 1000, 1500, 2000, 2500, 3000, 3500, 4000
L (m)	1000
α	0.8

R: radio range of the nodes, n: number of the nodes in the network, R_w: radio range of the colluding node, D: distance between the malicious node and the colluding node, L: side of the sensing area, α: critical possibility

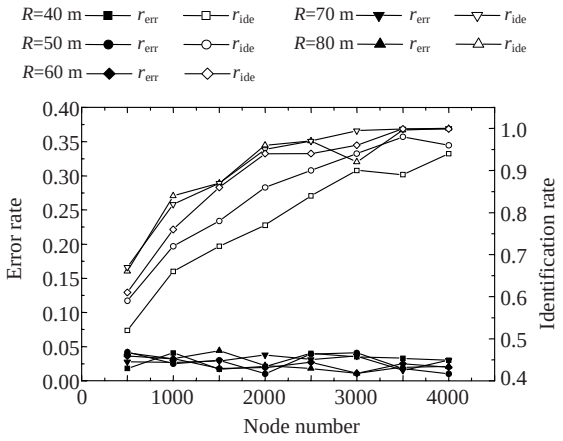

Fig.5 Error rate r_{err} and identification rate r_{ide} against the node number n

D, R_w and α are constant: D=500 m, R_w=60 m and α=0.8. The five curves relate to different radio ranges R of the sensor node

As α represents the scale of r_{err}, r_{err} should remain stable if α is constant. Fig.5 shows that r_{err} does not vary much with n and R. It fluctuates between 1% and 5%. In Fig.5, we find that r_{ide} rises as n and R increase. For n>2000 and R=50, 60, 70, 80 m, r_{ide} is greater than 80%. For n>3000 and R=70 m, r_{ide} even approximates to 1. The increase in r_{ide} with the increase in n and R is accentuated when the values of n and R are relatively small.

The interdependence of r_{ide} and T can account for this. To be more precise, r_{ide} relies on the ratio of $T/(np)$, not the absolute value of T. np is the mathematical expectation of a Bernoulli distribution like Eq.(6) and is also the average neighboring-node-number of a sensor node. Therefore, $T/(np)$ represents the degree to which the detection threshold exceeds the normal condition. So if $T/(np)$ is large, only the poisoned nodes whose neighboring-node-number significantly exceeds the average number can self-identify, and r_{ide} will be small. The opposite circumstance will happen if $T/(np)$ is small. Table 2 shows the values of $T/(np)$ with different n and R. It is obvious that $T/(np)$ decreases with the increase of n and R, so r_{ide} increases with n and R, as shown in Fig.5.

Generally, as the node number increases, the performance of r_{err} and r_{ide} remains stable or is even reinforced. This is a significant advantage of WAPN.

Table 2 $T/(np)$ with different n and R

n	$T/(np)$				
	R=40 m	50 m	60 m	70 m	80 m
500	1.99	2.04	1.79	1.69	1.59
1000	1.79	1.78	1.59	1.49	1.44
1500	1.72	1.61	1.53	1.39	1.36
2000	1.59	1.53	1.46	1.40	1.34
2500	1.59	1.48	1.38	1.35	1.31
3000	1.46	1.44	1.41	1.32	1.29
3500	1.48	1.42	1.36	1.32	1.26
4000	1.44	1.37	1.33	1.30	1.26

Performance of WAPN under different wormhole parameters, D and R_w

Fig.6 shows the performance of r_{err} and r_{ide} against R_w under the condition that n=2000, R=60 m, D=300 m, and α=0.8. We find that r_{err} remains stable or slightly decreases with an increase in R_w. Once the nodes' radio range and the positions of the sensor nodes and the wormhole are fixed, the nodes that misjudge themselves as poisoned nodes are also determined. Then, enlarging R_w can lead some misjudging nodes to be covered by the influencing area, which means that these misjudging nodes turn into real poisoned nodes. Consequently, r_{err} remains stable or decreases with R_w. However, compared with the sensing area, the enlargement of the influencing area is small. So the effect of enlarging R_w on r_{err} is slight. In Fig.6, for R_w=20, 40, 60, 80, and 100 m, r_{err} is 1.34%, 1.34%, 1.34%, 1.28%, and 1.23%, respectively. Similarly, enlarging R_w can result in more self-identifiable poisoned nodes and non-self-identifiable poisoned nodes being included by the influencing area. As a result r_{ide} fluctuates slightly with R_w, as shown in Fig.6.

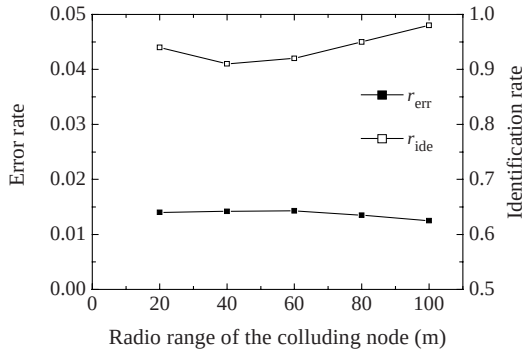

Fig.6 Error rate r_{err} and identification rate r_{ide} against radio range of the colluding node, R_w
n, D, R and α are constant: n=2000, D=500 m, R=60 m and α=0.8

Fig.7 shows r_{err} and r_{ide} against different D under the condition that n=2000, R=60 m, R_w=60 m and α=0.8. r_{err} is almost stable against different values of D. The uniform distribution of the sensor nodes in the sensing area can account for this stable performance of r_{err}. r_{ide} also remains stable (above 90%) for D=300, 500, 700, and 900 m. Fig.7 also shows that r_{ide} is relatively low for D=100 and 1300 m, especially for D=1300 m. When D is too small (D=100 m$<R_w+2R$), the radio range of some poisoned nodes may overlap with the radio range of the malicious node. So the nodes in the overlapping area are not recognized as additional neighbors of the poisoned nodes. For example, in Fig.8 the shaded area is the overlapping part of the radio ranges of the poisoned node N0 and the malicious node. In this case, the neighboring-node-number increase of some poisoned nodes (N0 for example) is reduced, so r_{ide} decreases. When D is too large (D=1300 m>L), the colluding node resides more

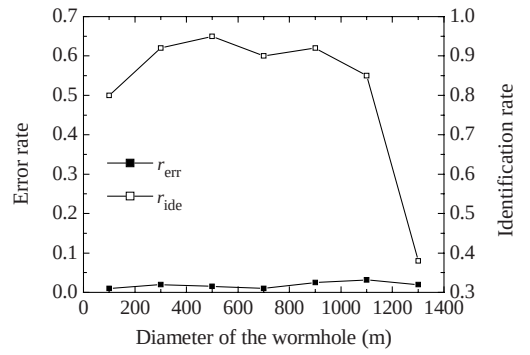

Fig.7 Error rate r_{err} and identification rate r_{ide} against diameter of the wormhole, D
n, R_w, R and α are constant: n=2000, R_w=60 m, R=60 m, and α=0.8

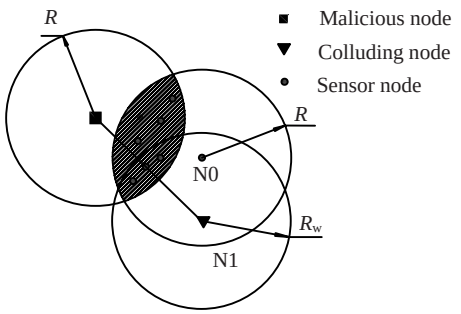

Fig.8 The overlapping case when D is small

often in the peripheral part of the sensing area (Fig.3). As a result, some poisoned nodes potentially have fewer neighbors than the average level. In this case, the increase of the neighboring-node-number resulting from the wormhole attack is also reduced. As a result r_{ide} again decreases. However, these two cases do not seriously affect the performance of WAPN, because the hazard caused by the wormhole in these two cases is very trivial. When the overlapping case (D is small) occurs, the two areas (the radio ranges of the malicious node and the colluding node) can be connected by the nodes in the overlapping area. So the main damage caused by the wormhole, traffic jamming, and mis-direction of the packets, is weakened. When the peripheral case (D is great) occurs, the colluding node already resides in the peripheral part. Few nodes take the poisoned nodes as their next-hop neighbors, so extension of mis-selection of the next-hop is restricted.

Effect of critical possibility on the performance of WAPN

The critical possibility, α, has a great influence on the performance of WAPN. As shown in Eq.(5), T is strongly related to α. At the same time, T is a dominant factor acting on r_{err} and r_{ide} and determines the tradeoffs between them. As T increases, fewer nodes self-identify themselves as poisoned nodes. So r_{ide} declines and r_{err} increases. When T decreases, more nodes self-identify themselves as poisoned nodes. As a result r_{ide} goes up and r_{err} declines. The change in r_{err} and r_{ide} against α under the conditions of different n and R is shown in Fig.9 and Fig.10, respectively. We find that a higher α can result in lower r_{err} and lower r_{ide}. Conversely, a smaller α can result in higher r_{err} and higher r_{ide}.

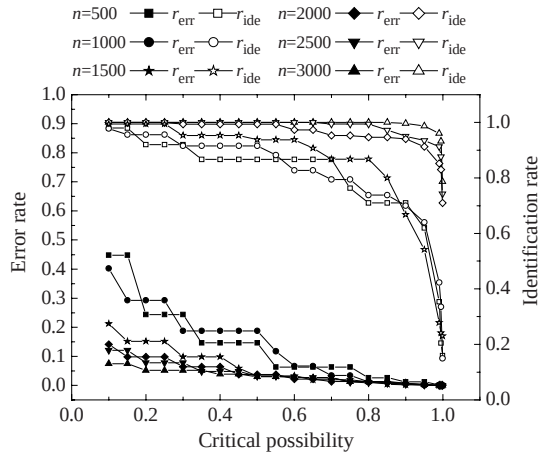

Fig.9 Error rate r_{err} and identification rate r_{ide} against critical possibility α at different sensor node numbers, n
D, R, and R_w are constant: D=500 m, R=60 m and R_w=60 m

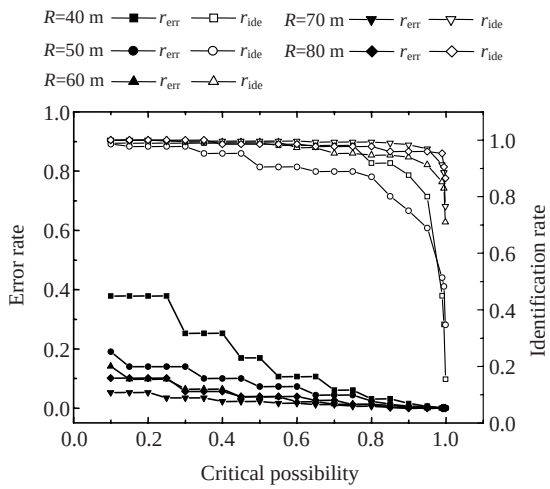

Fig.10 Error rate r_{err} and identification rate r_{ide} against critical possibility α at different radio ranges, R, of the sensor node
D, n and R_w are constant: D=500 m, n=2000 and R_w=60 m

CONCLUSION

This paper presents WAPN, a novel wormhole detection approach for WSNs based on the probability distribution of the neighboring-node-number. It differs from previous approaches that require extra hardware and functions of the sensor nodes or high computing and memory capacity of the sink. WAPN can be easily and distributively implemented. The only requirement is that each sensor node can obtain its neighboring-node-number, which can be easily achieved by most routing protocols. Simulations were carried out under the conditions of various network topologies and wormhole parameters. The results show that WAPN can detect the wormhole and the nodes influenced by the wormhole with a high identification rate and a low error rate in most cases. An additional advantage of WAPN is that its performance improves as node redundancy increases.

References

Aboelaze, M., Aloul, F., 2005. Current and Future Trends in Sensor Networks: A Survey. Proc. Int. Conf. on Wireless and Optical Communications Networks, Dubai, United Arab Emirates, p.551-555. [doi:10.1109/WOCN.2005.1436087]

Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E., 2002. Wireless sensor networks: a survey. *Computer Networks*, **38**(4):393-422. [doi:10.1016/S1389-1286(01)00302-4]

Boudriga, N., Obaidat, M.S., 2006. Mobility, sensing, and security management in wireless ad hoc sensor systems. *Comput. Electr. Eng.*, **32**(1-3):266-276. [doi:10.1016/j.compeleceng.2006.01.019]

Buttyan, L., Dora, L., Vajda, I., 2005. Statistical wormhole detection in sensor networks. *LNCS*, **3813**:128-141. [doi:10.1007/11601494_11]

Chong, C.Y., Kumar, S.P., 2003. Sensor networks: evolution, opportunities, and challenges. *Proc. IEEE*, **91**(8):1247-1265. [doi:10.1109/JPROC.2003.814918]

Doyle, B., Bell, S., Smeaton, A.F., McCusker, K., O′Connor, N.E., 2006. Security considerations and key negotiation techniques for power constrained sensor networks. *Comput. J.*, **49**(4):443-453. [doi:10.1093/comjnl/bxl023]

Faza, A.Z., Sedigh, S., 2006. A General Purpose Framework for Wireless Sensor Network Applications. Proc. 30th Annual Int. Computer Software and Applications Conf., Chicago, USA, p.356-358. [doi:10.1109/COMPSAC.2006.97]

Gura, N., Patel, A., Wander, A., Eberle, H., Chang Shantz, S., 2004. Comparing elliptic curve cryptography and RSA on 8-bit CPUs. *LNCS*, **3156**:925-943. [doi:10.1007/b99451]

Hu, L., Evans, D., 2004. Using Directional Antennas to Prevent Wormhole Attacks. Proc. 11th Network and Distributed System Security Symp., San Diego, USA, p.22-32.

Hu, Y.C., Perrig, A., Johnson, D.B., 2003. Packet Leashes: A Defense against Wormhole Attacks in Wireless Ad Hoc Networks. Proc. IEEE INFOCOM, San Francisco, CA, USA, **3**:1976-1986.

Karlof, C., Wagner, D., 2003. Secure routing in wireless sensor networks: attacks and countermeasures. *Ad Hoc Networks*, **1**(2-3):293-315. [doi:10.1016/S1570-8705(03)00008-8]

Lee, S.J., Gerla, M., 2001. Split Multipath Routing with Maximally Disjoint Paths in Ad Hoc Networks. Proc. IEEE Int. Conf. on Communications, Helsinki, Finland,

p.3201-3205. [doi:10.1109/ICC.2001.937262]

Poovendran, R., Lazos, L., 2007. A graph theoretic framework for preventing the wormhole attack in wireless ad hoc networks. *Wireless Networks*, **13**(1):27-59. [doi:10.1007/s11276-006-3723-x]

Qi, H., Iyengar, S.S., Chakrabarty, K., 2001. Distributed sensor networks a review of recent research. *J. Franklin Inst.*, **338**(6):655-668. [doi:10.1016/S0016-0032(01)00026-6]

Qian, L.J., Song, N., Li, X.F., 2007. Detection of wormhole attacks in multipath routed wireless ad hoc networks: a statistical analysis approach. *J. Network Comput. Appl.*, **30**(1):308-330. [doi:10.1016/j.jnca.2005.07.003]

Roman, R., Zhou, J., Lopez, J., 2005. On the security of wireless sensor networks. *LNCS*, **3482**(3):681-690. [doi:10.1007/b136271]

Shi, E., Perrig, A., 2004. Designing secure sensor networks. *IEEE Wirel. Commun.*, **11**(6):38-43. [doi:10.1109/MWC.2004.1368895]

Slijepcevic, S., Potkonjak, M., Tsiatsis, V., Zimbeck, S., Srivastava, M.B., 2002. On Communication Security in Wireless Ad-hoc Sensor Networks. Proc. 11th IEEE Int. Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, p.139-144. [doi:10.1109/ENABL.2002.1030000]

Sohrabi, K., Gao, J., Ailawadhi, V., Pottie, G.J., 2000. Protocols for self-organization of a wireless sensor network. *IEEE Pers. Commun.*, **7**(5):16-27. [doi:10.1109/98.878532]

Wang, W.C., Bhargava, B., 2004. Visualization of Wormholes in Sensor Networks. Proc. ACM Workshop on Wireless Security, Philadelphia, PA, USA, p.51-60.

APPENDIX

Table A1 Threshold T against α and np

α	T							
	np=10	20	30	40	50	60	70	80
0.10	13	26	38	50	61	73	84	95
0.15	13	26	38	50	62	73	84	96
0.20	14	26	38	50	62	73	85	96
0.25	14	26	38	50	62	74	85	96
0.30	14	27	39	51	62	74	85	97
0.35	14	27	39	51	63	74	86	97
0.40	14	27	39	51	63	75	86	98
0.45	15	27	40	52	63	75	87	98
0.50	15	28	40	52	64	75	87	98
0.55	15	28	40	52	64	76	87	99
0.60	15	28	41	53	65	76	88	99
0.65	16	29	41	53	65	77	89	100
0.70	16	29	42	54	66	78	89	101
0.75	16	30	42	54	66	78	90	102
0.80	17	30	43	55	67	79	91	102
0.85	17	31	44	56	68	80	92	104
0.90	18	32	45	57	69	81	93	105
0.95	19	33	46	59	71	84	96	108

α: critical possibility; n: number of the nodes in the network; p: probability of any two nodes being neighbors

Table A2 $F(T)$ against α and np

α	$F(T)$							
	np=10	20	30	40	50	60	70	80
0.10	0.86	0.92	0.94	0.95	0.94	0.96	0.96	0.96
0.15	0.86	0.92	0.94	0.95	0.96	0.96	0.96	0.96
0.20	0.92	0.92	0.94	0.95	0.96	0.96	0.96	0.96
0.25	0.92	0.92	0.94	0.95	0.96	0.97	0.96	0.96
0.30	0.92	0.95	0.95	0.96	0.96	0.97	0.96	0.97
0.35	0.92	0.95	0.95	0.96	0.97	0.97	0.97	0.97
0.40	0.92	0.95	0.95	0.96	0.97	0.97	0.97	0.98
0.45	0.95	0.95	0.97	0.97	0.97	0.97	0.98	0.98
0.50	0.95	0.97	0.97	0.97	0.98	0.97	0.98	0.98
0.55	0.95	0.97	0.97	0.97	0.98	0.98	0.98	0.98
0.60	0.95	0.97	0.98	0.98	0.98	0.98	0.98	0.98
0.65	0.97	0.98	0.98	0.98	0.98	0.99	0.99	0.99
0.70	0.97	0.98	0.99	0.99	0.99	0.99	0.99	0.99
0.75	0.97	0.99	0.99	0.99	0.99	0.99	0.99	0.99
0.80	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
0.85	0.99	0.99	0.99	0.99	0.99	0.99	1.00	1.00
0.90	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

$F(T)$: probability of a certain node having less than T neighboring nodes