ActiveTrust : Secure and Trustable Routing in Wireless Sensor Networks

メタデータ	言語: eng 出版者: IEEE 公開日: 2016-10-18 キーワード (Ja): キーワード (En): wireless sensor networks, black hole attack, network lifetime, security, trust 作成者: LIU, Yuxin, 董, 冕雄, 太田, 香, LIU, Anfeng メールアドレス: 所属:
URL	http://hdl.handle.net/10258/00009018

ActiveTrust: Secure and Trustable Routing in Wireless Sensor Networks

Yuxin Liu, Mianxiong Dong, *Member, IEEE,* Kaoru Ota, *Member, IEEE,* Anfeng Liu

***Abstract*— Wireless sensor networks (WSNs) are increasingly being deployed in security-critical applications. Because of their inherent resource-constrained characteristics, they are prone to various security attacks, and a black hole attack is a type of attack that seriously affects data collection. To conquer that challenge, an active detection-based security and trust routing scheme named ActiveTrust is proposed for WSNs. The most important innovation of ActiveTrust is that it avoids black holes through the active creation of a number of detection routes to quickly detect and obtain nodal trust and thus improve the data route security. More importantly, the generation and distribution of detection routes are given in the ActiveTrust scheme, which can fully use the energy in non-hotspots to create as many detection routes as needed to achieve the desired security and energy efficiency. Both comprehensive theoretical analysis and experimental results indicate that the performance of the ActiveTrust scheme is better than that of previous studies. ActiveTrust can significantly improve the data route success probability and ability against black hole attacks and can optimize network lifetime.**

***Index Terms*—black hole attack, network lifetime, security, trust, wireless sensor networks**

I. INTRODUCTION

WIRELESS Sensor Networks (WSNs) are emerging as a promising technology because of their wide range of applications in industrial, environmental monitoring, military and civilian domains [1-5]. Due to economic considerations, the nodes are usually simple and low cost. They are often unattended, however, and are hence likely to suffer from different types of novel attacks [6-8]. A black hole attack (BLA) is one of the most typical attacks [9] and works as follows. The adversary compromises a node and drops all packets that are routed via this node, resulting in sensitive data being discarded or unable to be forwarded to the sink. Because the network makes decisions depending on the nodes' sensed data, the consequence is that the network will completely fail and, more seriously, make incorrect decisions [10-15]. Therefore, how to detect and avoid BLA is of great significance for security in WSNs.

There is much research on black hole attacks [9, 16-19]. Such studies mainly focus on the strategy of avoiding black holes [17, 18, 19]. Another approach does not require black hole information in advance. In this approach, the packet is divided into *M* shares, which are sent to the sink via different routes (multi-path), but the packet can be resumed with *T* shares (*T*<=*M*). However, a deficiency is that the sink may receive more than the required *T* shares, thus leading to high energy consumption; such research can be seen in [9, 16]. Another preferred strategy that can improve route success probability is the trust route strategy. There is much related research, such as [20, 21, 22, 23, 24]. The main feature is to create a route by selecting nodes with high trust because such nodes have a higher probability of routing successfully; thus, routes created in this manner can forward data to the sink with a higher success probability [22, 23].

However, the current trust-based route strategies face some challenging issues [24]. (1) The core of a trust route lies in obtaining trust. However, obtaining the trust of a node is very difficult, and how it can be done is still unclear. (2) Energy efficiency. Because energy is very limited in WSNs, in most research, the trust acquisition and diffusion have high energy consumption, which seriously affects the network lifetime. (3) Security. Because it is difficult to locate malicious nodes, the security route is still a challenging issue. Thus, there are still issues worthy of further study. Security and trust routing through an active detection route protocol is proposed in this paper. The main innovations are as follows.

(1) The ActiveTrust scheme is the first routing scheme that uses active detection routing to address BLA.

The most significant difference between ActiveTrust and previous research is that we create multiple detection routes in regions with residue energy; because the attacker is not aware of detection routes, it will attack these routes and, in so doing, be exposed. In this way, the attacker's behavior and location, as well as nodal trust, can be obtained and used to avoid black holes when processing real data routes. To the best of our knowledge, this is the first proposed active detection mechanism in WSNs.

Manuscript received 17 Dec 2015; revised February 6, 2016; accepted May 5, 2016. Date of publication **** **, 2016; date of current version May 9, 2016.

This work was supported in part by the National Natural Science Foundation of China (61379110, 61073104, 61572528, 61272494, 61572526), The National Basic Research Program of China (973 Program) (2014CB046305), JSPS KAKENHI Grant Number 26730056, 16K00117, 15K15976, and JSPS A3 Foresight Program. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. *****. (Corresponding author: Anfeng Liu).

Yuxin Liu, Anfeng Liu are with School of Information Science and Engineering, Central South University, Changsha, 410083 China. E-mail: yuxinliu@csu.edu.cn, afengliu@mail.csu.edu.cn.

Mianxiong Dong and Kaoru Ota are with Muroran Institute of Technology, Japan. E-mail: {mx.dong, ota}@csse.muroran-it.ac.jp.

(2) The ActiveTrust route protocol has better energy efficiency.

Energy is very precious in WSNs, and there will be more energy consumption if active detection is processed. Therefore, in previous research, it was impossible to imagine adopting such high-energy-consumption active detection routes. However, we find it possible after carefully analyzing the energy consumption in WSNs. Research has noted that there is still up to 90% residue energy in WSNs when the network has died due to the "energy hole" phenomenon. Therefore, the ActiveTrust scheme takes full advantage of the residue energy to create detection routes and attempts to decrease energy consumption in hotspots (to improve network lifetime). Those detection routes can detect the nodal trust without decreasing lifetime and thus improve the network security. According to theoretical analysis and experimental results, the energy efficiency of the ActiveTrust scheme is improved more than 2 times compared to previous routing schemes, including shortest routing, multi-path routing.

(3) The ActiveTrust scheme has better security performance. Compared with previous research, nodal trust can be obtained in ActiveTrust. The route is created by the following principle. First, choose nodes with high trust to avoid potential attack, and then route along a successful detection route. Through the above approach, the network security can be improved.

(4) Through our extensive theoretical analysis and simulation study, the ActiveTrust routing scheme proposed in this paper can improve the success routing probability by 1.5 times to 6 times and the energy efficiency by more than 2 times compared with that of previous researches.

The rest of this paper is organized as follows. In Section 2, the related work is reviewed. The system model and problem statement are described in Section 3. In Section 4, the novel ActiveTrust scheme is presented. Security and performance analyses are provided in Section 5. Section 6 is the analysis and comparison of experimental results. We conclude in Section 7.

II. Related Work

Single-path routing is a simple routing protocol [12] but is easily blocked by the attacker. Therefore, the most natural approach is via multi-path routing to the sink. Even if there is an attack in some route, the data can still safely reach the sink [9]. Multi-path routing protocols can be classified into two classes depending on whether the data packet is divided. One is multi-path routing without share division. The other is multi-path routing with share division, i.e., the packet is divided into shares, and different shares reach the destination via different routes [9].

(1) Non-share-based multi-path routing. There are different multi-path route construction methods. Ref. [25] proposes a multi dataflow topologies (MDT) approach to resist the selective forwarding attack. In the MDT approach, the network is divided into two dataflow topologies. Even if one topology has a malicious node, the sink can still obtain packets from the other topology.

In such protocols, the deficiency is that if the packet is routed via *n* routes simultaneously, the energy consumption will be *n* times that of a single path route, which will seriously affect the network lifetime; similar research can be seen in multi-path DSR [25], the AOMDV [18] and AODMV [26].

(2) Share-based multi-path routing protocols. The SPREAD algorithm in [27] is a typical share-based multi-path routing protocol. The basic idea of the SPREAD algorithm is to transform a secret message into multiple shares, which is called a (T, M) threshold secret sharing scheme [28]. The M shares are delivered by multiple independent paths to the sink such that, even if a small number of shares are dropped, the secret message as a whole can still be recovered [9, 16, 28]. The advantage of this algorithm is that through multi-path routing, each path routes only one share, and the attacker must capture at least T shares to restore nodal information, which increases the attack difficulty [9]. Thus, the privacy and security can be improved. In the above research, the multi-path routing algorithms are deterministic such that the set of route paths is predefined under the same network topology [9]. This weakness opens the door for various attacks if the routing algorithm is obtained by the adversary [9].

For the weakness mentioned above, Ref. [29] proposed four random propagation strategies: random propagation (PRP), directed random propagation (DRP), non-repetitive random propagation (NRRP), and multicast tree assisted random propagation (MTRP). The general strategy is as follows. First, divide the message into M shares, and the route path of each share is not predetermined. Thus, even if the adversary acquires the routing algorithm, it is difficult to launch a pinpointed node-compromise or jamming attack. Because it is difficult to capture more than T shares, the security is also improved. In multi-to-one data collection WSNs, we argue that for classic "slicing and assembling" or multi-path routing techniques, sliced shares will merge in the same path with high probability, and this path can be easily attacked by black holes. Thus, in [16], a Security- and Energy-efficient Disjoint Route (SEDR) scheme is proposed to route sliced shares to the sink with randomized disjoint multipath routes by utilizing the available surplus energy of sensor nodes. The authors demonstrate that the security is maximized without reducing the lifetime in the SEDR protocol.

Another method to avoid attack and improve route success probability is trust routing. Trust management [20] is becoming a new driving force for solving challenges in ad hoc networks [21], peer-to-peer networks [22], and WSNs [23, 24].

Zhan et al proposed a trust-aware routing framework protocol (TARF), using trust and energy cost for route decisions, to prevent malicious nodes from misleading network traffic [30]. Ref. [31] proposes the Sec-CBSN algorithm, which develops different trust calculation methods based on nodal roles. Ref. [32] develops an attack-resistant and lightweight trust management protocol named ReTrust, which can resist attacks through a trust management approach for medical sensor networks (MSNs). Ref. [33] presents a proposal named TRIP, which aims to quickly and accurately identify malicious or selfish nodes spreading false information in vehicular ad hoc networks (VANETs). Ref. [34] also proposes a resilient trust model, SensorTrust, for hierarchical WSNs. Ref. [24] introduces the concept of attribute similarity in finding potentially friendly nodes among strangers.

Although there is much research on black node attack avoidance, there is still much that is worthy of further study. (1) The current black hole avoidance strategies mostly affect network lifetime. (2) The current black hole avoidance strategies are mostly passive acting systems, which affects system performance. (3) The trust route mechanism has high costs and is difficult to obtain trust, so the guiding significance is limited [35, 36].

III. The System Model and Problem Statement

A. The System Model

(1) Network model

(a) We consider a wireless sensor network consisting of sensor nodes that are uniformly and randomly scattered in a circular network; the network radius is R, with nodal density ρ, and nodes do not move after being deployed [4, 9]. Upon detection of an event, a sensor node will generate messages, and those messages must be sent to the sink node [4, 13].

(b) We consider that link-level security has been established through a common cryptography-based protocol. Thus, we consider a link key to be safe unless the adversary physically compromises either side of the link [9, 16].

(2) The adversaries model

We consider that black holes are formed by the compromised nodes and will unselectively discard all packets passed by to prevent data from being sent to the sink [9, 16]. The adversary has the ability to compromise some of the nodes. However, we consider the adversary to be unable to compromise the sink and its neighboring nodes [9, 16].

B. Energy Consumption Model and Related Definitions

According to the typical energy consumption model [4, 9, 16], Eq. (1) represents energy consumption for transmitting, and Eq. (2) represents energy consumption for receiving.

E_{elec} represents the transmitting circuit loss. Both the free space (d^2 power loss) and the multi-path fading (d^4 power loss) channel models are used in the model depending on the distance between the transmitter and receiver. ε_{fs} and ε_{amp} are the respective energy required by power amplification in the two models. The energy consumption for receiving an l-bit packet is shown in Eq. (2). The above parameter settings are shown in Table 1, as adopted from [4, 9, 16].

$$\begin{cases} E_{member} = lE_{elec} + l\varepsilon_{fs}d^2 & if\ d \le d_0 \\ E_{member} = lE_{elec} + l\varepsilon_{amp}d^4 & if\ d > d_0 \end{cases} \quad (1)$$

$$E_R(l) = lE_{elec} \quad (2)$$

Table 1 network parameters

Parameter	Value
Threshold distance (d_0) (m)	87
Sensing range r_s (m)	15
E_{elec} (nJ/bit)	50
e_{fs} (pJ/bit/m^2)	10
e_{amp} (pJ/bit/m^4)	0.0013
Initial energy (J)	0.5

C. Problem Statement

(1) Network lifetime maximization. Network lifetime can be defined as the first node die time in the network [4, 9, 16]. For E_i as the energy consumption for node i, the lifetime maximization can be expressed as the following:

$$\max(T) = \min\ \max(E_i) \quad (3)$$

(2) The data collection has better security performance and strong capability against black hole attacks.

The main goal of our scheme is to ensure that the nodal data safely reach the sink and are not blocked by the black hole. Thus, the scheme design goal is to maximize the ratio of packets successfully reaching the sink. Consider that the number of packets that are required to reach the sink is M and that the number of packets that ultimately succeed in reaching the sink is m; the success ratio is

$$q = m/\mathrm{M} \quad (4)$$

Our goal is to maximize the success ratio, that is, $\max(q)$. In summary, the optimization goal of this paper is the following equation:

$$\begin{cases} \max(T) = \min\ \max\limits_{0<i\le n}(E_i) \\ \max(q) \mid q = m/\mathrm{M} \end{cases} \quad (5)$$

IV. Active Trust Scheme Design

A. Overview of the Proposed Scheme

An overview of the ActiveTrust scheme, which is composed of an active detection routing protocol and data routing protocol, is shown in Fig. 1.

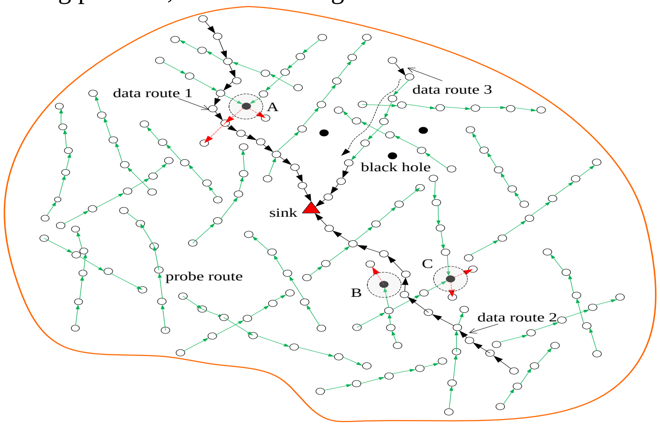


Fig. 1: Illustration of the ActiveTrust scheme

(1) Active detection routing protocol: A detection route refers to a route without data packets whose goal is to convince the adversary to launch an attack so the system can identify the attack behavior and then mark the black hole location. Thus, the system can lower the trust of suspicious nodes and increment the trust of nodes in successful routing routes. Through active detection routing, nodal trust can be quickly obtained, and it can effectively guide the data route in choosing nodes with high trust to avoid black holes. The active detection routing protocol is shown via the green arrow in Fig. 1. In this scheme, the source node randomly selects an undetected neighbor node to create an active detection route. Considering that the longest detection route length is ϖ, the detection route decreases its

length by 1 for every hop until the length is decreased to 0, and then the detection route ends.

(2) Data routing protocol. The data routing refers to the process of nodal data routing to the sink. The routing protocol is similar to common routing protocols in WSNs [3, 7, 8]; the difference is that the route will select a node with high trust for the next hop to avoid black holes and thus improve the success ratio of reaching the sink.

The data routing is shown via the black arrow in Fig. 1. The routing protocol can adopt an existing routing protocol [7, 12], and we take the shortest route protocol as an example. Node a in the route will choose the neighbor that is nearer the sink and has high trust as the next hop. If there is not a node among all neighbors nearer the sink that has trust above the default threshold, it will report to the upper node that there is no path from a to the sink. The upper node, working in the same manner, will re-select a different node from among its neighbors nearer the sink until the data are routed to the sink or there is conclusively no path to the sink.

B. *Active Detection Routing Protocol*

Table 2: Pseudo-code of Algorithm 1 for the active detection routing protocol

Algorithm 1: Active Detection Routing Protocol

```
Initialization
For each neighbor node A_n Do
    Let A_n.accesTime=Current_time
End for
For each node that generates a detection packet, such as node A, Do
    Construct packet P, and do value assignment for $\omega$ and $\varpi$
    Select B as the next hop which B meets access time is the minimum and
     nearer the sink
    //B is the node that is the longest time undetected and nearer the sink
    Send packet P to node B
End for
For each node that receives a detection packet, such as node B, Do
    let P.$\omega$ =P.$\omega$ -1, P.$\varpi$ =P.$\varpi$ -1
    If $\varpi$ =0 then
        Construct feedback packet q, and do value assignment for each part
        Send feedback packet q to the source
    End if
    If p.$\omega \neq$ 0 then
        detection routing continue
    End if
End for
For each node that receives feedback packet q, such as node C, Do
    If q.destination is not itself then
        send q to the source node
    End if
End for
```

This section details the implementation of the active detection routing protocol. The content of the detection routing packet can be divided into 6 parts, as shown in Fig. 2: (a) packet head; (b) packet type; (c) ID of the source node; (d) maximum detection route length; (e) acknowledge returned to the source for every ω hops; and (f) ID of the packet.

head	type	source	ϖ	ω	id

Fig. 2: The structure of packets of detection routes

The source node selects an undetected node to launch the detection route. Once the detection packet is received by nodes, the maximum route length ϖ is decreased by 1. After that, if ϖ is 0, generate a feedback packet and launch a feedback route to the source, and then restore ϖ to the initial value. If ϖ is not 0, then continue to select the next hop in the same way; otherwise, end the route. The structure of a feedback packet is shown in Fig. 3, and it is also composed of 6 parts: (a) packet head; (b) packet type; (c) ID of the source node; (d) destination node; (e) ID of the detection packet; and (f) ID of the packet.

head	type	source	Destination	S-id	id

Fig. 3: The structure of feedback packets of a detection route

The feedback packet is routed back to the data source; because nodes cache the detection route info, the feedback packet is able to return back to the source, and the following is the algorithm for the detection route protocol.

C. *Calculation of Nodal Trust*

During data routing and detection routing, every node will perform a nodal trust calculation to aid in black hole avoidance. When node A performs a detection route for node B at time t_i, if the detection data are successfully routed, consider the trust of node A to B to be $\Delta_A^B(t_i)$; otherwise, consider the trust to be $\Lambda_A^B(t_i)$. Considering that A has w interactions with B during t, the detection value order by time is as follows:

$$\left\{\Delta_A^B(t_1) \mid \Lambda_A^B(t_1),\ \Delta_A^B(t_2) \mid \Lambda_A^B(t_2),\ \ldots\ \Delta_A^B(t_w) \mid \Lambda_A^B(t_w)\right\}$$

$\Delta_A^B(t_i) \mid \Lambda_A^B(t_i)$ refers to the trust value of A to B at t_i (if data are dropped, then $\Lambda_A^B(t_i)$ <0; otherwise, $\Delta_A^B(t_i)$ >0).

Definition 1 (**Nodal direction trust**): Consider the trust set of node A to node B during t to be:

$$\left\{\Delta_A^B(t_1) \mid \Lambda_A^B(t_1),\ \Delta_A^B(t_2) \mid \Lambda_A^B(t_2),\ \ldots\ \Delta_A^B(t_w) \mid \Lambda_A^B(t_w)\right\}$$

Then, during period t, the total direction trust of A to B is:

$$\begin{cases} C_A^B = \sum_{i=1}^{w}\left\{\left(\Delta_A^B(t_i) \mid \Lambda_A^B(t_i)\right)\cdot \hbar(i)\right\} \Big/ w, & w \neq 0 \\ 0, & w = 0 \end{cases} \quad (6)$$

In Eq. (6), $\hbar(i) \in [0,1]$ is an attenuation function to weight direction trusts at different times; according to common sense, the latest behavior should be given more weight [24], and otherwise less weight. The attenuation function is as shown in Eq. (7), and γ is a decimal less than 1.

$$\hbar(i) = \begin{cases} 1, & i = w \\ \hbar(i-1) = \gamma\hbar(i), & 1 \le i < w \end{cases} \quad (7)$$

In the ActiveTrust scheme, the trust calculation should meet the following condition. If the node is found to be malicious in the latest detection, then its trust should be below the threshold Θ, and the node will not be chosen for later routing. If the malicious node returns to the normal node, it needs several detections to take it into routing consideration; thus, the parameter γ should meet the following equation:

Theorem 1: Consider that there are at most w interactions involved in the trust computation and that the threshold is Θ; then, the parameter γ should meet the following equation:

$$\left(\left(1-\gamma^{w}\right)/\left(1-\gamma\right)-1\right)\cdot\Delta_A^B<\Theta-\Lambda_A^B \qquad (8)$$

Proof: If the node is shown to be malicious in the latest detection, then we can obtain $\Lambda_A^B(t_i)$ <0; if it was shown to be trustable in the previous $w-1$ detections, then the trust of node A to node B must meet the following formula:

$$\Lambda_A^B+\gamma\Delta_A^B+\gamma^2\Delta_A^B+\ \ldots+\gamma^{w-1}\Delta_A^B<\Theta$$

$$\Rightarrow\left(\left(1-\gamma^{w}\right)/\left(1-\gamma\right)-1\right)\cdot\Delta_A^B<\Theta-\Lambda_A^B$$

If there is more than one malicious result in the previous $w-1$ detections, the trust should be less than Θ, thus proved.

■

Inference 1: If the node is shown to be malicious, then when it returns to normal, there must be at least η trustable detections, and it can be re-considered a trustable node; η meets the following:

$$1+\gamma+\ldots+\gamma^{\eta-1}+\gamma^{\eta}\partial+\gamma^{\eta+1}+\ldots+\gamma^{w-1}>\Theta/\Delta_A^B \qquad (9)$$

Proof: Consider that a node is shown to be trustable in the current η detections, that is, Δ_A^B >0, and malicious in the later detection, namely, Λ_A^B<0. Additionally, in the previous $w-1$ detections, the behaviors were all trustable. In this situation, η is the minimum, and the trust of A to B at this time is as follows:

$$\Delta_A^B,\gamma\Delta_A^B,\ldots,\gamma^{\eta-1}\Delta_A^B,\gamma^{\eta}\Lambda_A^B,\gamma^{\eta+1}\Delta_A^B,\ldots,\gamma^{w-1}\Delta_A^B$$

The trust calculation is

$$\Delta_A^B+\gamma\Delta_A^B+\ldots+\gamma^{\eta-1}\Delta_A^B+\gamma^{\eta}\Lambda_A^B+\gamma^{\eta+1}\Delta_A^B+\ldots+\gamma^{w-1}\Delta_A^B>\Theta$$

Considering that $\Lambda_A^B(t_i)=\partial\Delta_A^B$, the above can be transformed into

$$1+\gamma+\ldots+\gamma^{\eta-1}+\gamma^{\eta}\partial+\gamma^{\eta+1}+\ldots+\gamma^{w-1}>\Theta/\Delta_A^B$$

■

Definition 2 (**Nodal recommendation trust**): Node A is the trust evaluator, node C is the target of evaluation, and node B is a recommender of A. Consider C_A^B to be the direction trust of A to B and C_B^C to be the direction trust of B to C; then, the recommendation trust of A to C is

$$R_A^C=C_A^B\times C_B^C \qquad (10)$$

For the trust of multiple recommendations, the calculation of the recommendation trust from A to B, B to C, etc., until D to E is

$$R_A^E=C_A^B\times C_B^C\times C_C^D\times C_D^E \qquad (11)$$

Definition 3 (Recommendation trust merging): Consider that the recommender set of node A is R_A, $n_i\in R_A$ and that the recommendation trust of n_i to node K is $R_A^{i,k}$; then, the merged trust of A to K is

$$U_A^K=\sum_{n_i\in A_n}\left(u_{n_i}R_A^{n_i,k}\right)\ |u_{n_i}=\frac{R_A^{n_i,k}}{\left(R_A^{n_1,k}+R_A^{n_2,k}+\ldots+R_A^{n_{m-1},k}+R_A^{n_m,k}\right)} \qquad (12)$$

Definition 4 (Comprehensive trust): Comprehensive trust is the total trust, which merges the recommendation trust and direction trust:

$$C_{A,B}^T=\delta C_A^B+\left(1-\delta\right)U_A^B \qquad (13)$$

The comprehensive trust of a node can be computed as follows. After the node launches a detection route, it calculates the direction trust according to Eq. (6) for each received feedback packet. Through interactions, the node obtains the recommendation trust from its neighbors according to Eq. (10), and it then calculates the merged trust according to Eq. (12) for the multiple-recommendation trust. Finally, it calculates the comprehensive trust according to Eq. (13).

D. *Data Routing Protocol*

The core idea of data routing is that when any node receives a data packet, it selects one node from the set of candidates nearer the sink whose trust is greater than the preset threshold as the next hop. If the node cannot find any such appropriate next hop node, it will send a feedback failure to the upper node, and the upper node will re-calculate the unselected node set and select the node with the largest trust as the next hop; similarly, if it cannot find any such appropriate next hop, it sends a feedback failure to its upper node. The protocol is as follows:

Table 3: Pseudo-code of Algorithm 2 for data routing protocol

Algorithm 2: Data Routing Protocol

```
1:  For each node that generates or receives a data packet, such as node A, Do
2:      select B as the next hop such that B has never been selected in this data routing process, has the largest trust and is nearer the sink
4:      If A finds such node, for instance, node B
5:          Send data packet P to node B
6:          If node B is the sink then
7:              this data routing procession is completed
8:          End if
9:      Else
10:         Send failure feedback to the upper node, such as node C
11:     End if
12: End for
13: For each node that receives failure feedback, such as node B, Do
14:     Repeat step 2 to step 11
15: End for
```

E. *The Number of Active Detection Routes*

First, we analyze the energy consumption at different distances from the sink. As in theorem 1 of Ref. [16], consider the network radius to be R, the nodal transmission radius to be r, and the event generation rate to be λ; the shortest route path protocol is deployed such that the nodal distance to the sink is l, $l=hr+x$. The number of data packets undertaken by this node is thus as follows:

$$d_l=\left(\left(z+1\right)+\left(z\left(1+z\right)r\right)/2l\right)\lambda \qquad (14)$$

$$|\ z \text{ is an integer that makes } l+zr \text{ just smaller than } R$$

From Eq. (14), the energy consumption depends on the undertaken data amount. Thus, this paper considers the nodal data amount to represent the nodal load. Because the network lifetime depends on the node that has the highest energy consumption, we consider the maximum nodal data load to be

$d_{\max}$ and the energy consumption to be $d_{\max}e_u$ and thus observe that there is remaining energy for nodes whose data load is smaller than $d_{\max}$; then, we can fully use the remaining energy to construct detection routes. For the node whose distance to the sink is l, the remaining energy of the node is $(d_{\max}-d_l)e_u$, which can be used for detection. If the distance of an active detection route is measured by hops, then the available nodal hops of the active detection route is as follows:

Theorem 2: If the nodal distance to the sink is l, then the maximum detection hops that can be achieved by its residue energy is $\hbar_l = \left((d_{\max}-d_l)(1+\kappa_2)\right)/(1+\kappa_2/\kappa_1)$, where κ_1 is the ratio of data packet length to detection packet length and κ_2 is the ratio of data packet length to head packet length.

Proof: According to Eq. (14), for a node whose distance to the sink is l, its data load is $d_l = \left((z+1)+\left(z(1+z)r\right)/2l\right)\lambda$. The maximum nodal data load is $d_{\max} = \left((z+1)+\left(z(1+z)r\right)/2l_{\min}\right)\lambda$. Thus, the residue energy of this node is $(d_{\max}-d_l)e_p$, where e_p denotes the energy consumption for sending and receiving a unit data packet. Considering that the energy consumption for sending and receiving one bit data is e_u, $e_p = \xi e_u$ because ξ is the unit packet length, $\xi = \xi_h + \xi_b$, ξ_h is the packet head length, and ξ_b is the packet body length. Then, the available residue energy is $(d_{\max}-d_l)e_u(\xi_h+\xi_b)$ because the energy consumption for sending and receiving one detection packet is $e_u(\zeta_h+\zeta_b)$, where ζ_h is the head packet length of the detection packet and ζ_b is the body packet length. Consider ξ_h to equal ζ_h, namely, $\xi_h = \zeta_h$, $\xi_b = \kappa_1\zeta_b$, $\xi_b = \kappa_2\xi_h$.

Then, the active detection route hops that can be achieved by the nodal residue energy is

$$\hbar_l = (d_{\max}-d_l)e_u(\xi_h+\xi_b)\Big/\left(e_u(\zeta_h+\zeta_b)\right)$$

$$\Rightarrow \hbar_l = \left((d_{\max}-d_l)(1+\kappa_2)\right)\Big/(1+\kappa_2/\kappa_1) \qquad \blacksquare$$

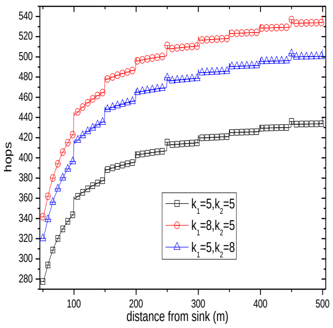


Fig. 4 The maximum detection hops afforded by the residue energy of nodes (different k_1, k_2)

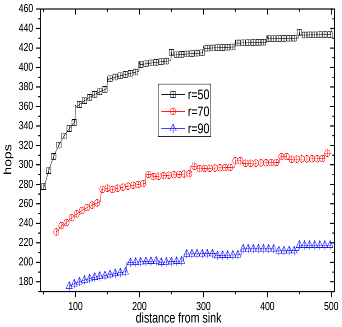


Fig. 5 The maximum detection hops afforded by the residue energy of nodes (different r)

Figs. 4 and 5 provide the maximum detection hops afforded by the residue energy of nodes with different distances from the sink. As seen, there is much residue energy in non-hotspots because the detection packet length is small; in a network with radius R =500 m, the detection hops can number in the hundreds, which shows that the network has sufficient energy to process detection routes without affecting the network lifetime.

Theorem 3: If the detection route length is ϖ hops and one detection feedback packet is returned to the detection source every ω ($\omega \le \varpi$) hops, then the total number of detection hops in this route is

$$\hbar_{\varpi,\omega} = \sum_{k=1}^{i} k\omega + 2\varpi \quad | i = \lfloor \varpi/\omega \rfloor \tag{15}$$

Proof: Because the detection data route length is ϖ hops, the number of data route hops is ϖ. One detection feedback is returned every ω hops for a route with length ϖ, and feedback is returned at $\omega, 2\omega, \dots i\omega, \varpi$, where $i\omega \le \varpi$. The number of hops for each returned feedback is $\omega, 2\omega, \dots i\omega, \varpi$; the number of returned packet hops is thus $\sum_{k=1}^{i} k\omega + \varpi$. Because the route length is ϖ and because it is possible for part of the route to be unable to be created or for returned packets to be unable to reach the detection source due to malicious nodes, the maximum number of detection hops is $\sum_{k=1}^{i} k\omega + 2\varpi$.

$\blacksquare$

In summary, the number of detection routes that can be created by residue energy can be found via Inference 2.

Inference 2: For a node whose distance to the sink is l, where the detection route length is ϖ hops and one detection feedback packet is returned to the detection source every ω ($\omega \le \varpi$) hops, the number of routes created by the residue energy is

$$\gamma_{\varpi,\omega} = \left(\frac{(d_{\max}-d_l)(1+\kappa_2)}{(1+\kappa_2/\kappa_1)}\right)\Bigg/\left(\sum_{k=1}^{i} k\omega + 2\varpi\right) \quad | i = \left\lfloor \frac{\varpi}{\omega} \right\rfloor \tag{16}$$

Proof: According to theorem 2, for a node at a distance l from the sink, the maximum detection hops that can be achieved by its residue energy is $\hbar_l = \left((d_{\max}-d_l)(1+\kappa_2)\right)/(1+\kappa_2/\kappa_1)$; theorem 3 shows the maximum number of detection hops to be $\sum_{k=1}^{i} k\omega + 2\varpi$. Thus, the number of detection routes can be obtained by dividing these two values. $\blacksquare$

Inference 3: The ActiveTrust scheme has the same network lifetime as do schemes without any security strategy.

Proof: The ActiveTrust scheme uses residue energy to construct detection routes; this construction energy consumption will not make the nodal energy consumption larger than $d_{\max}$, and thus, its lifetime is still $\Phi = E_{init}/d_{\max}$.

$\blacksquare$

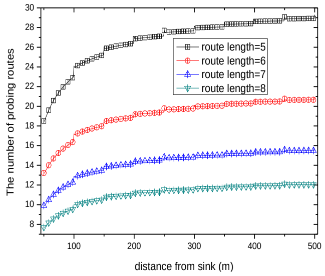


Fig. 6 The number of probing routes that can be created

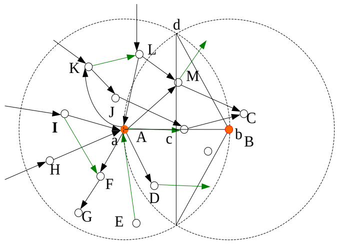


Fig. 7 Indirection trust among nodes

Fig. 6 shows, in a network with radius R =500 m, the number of probing routes that can be created by residue energy in non-hotspots under detection route lengths of 5, 6, 7, and 8. As seen, the residue energy can support at least 7 detection routes. Performance Analysis of the ActiveTrust Scheme

F. Analysis of the Successful Routing Probability

Theorem 3: Considering that the nodal degree is d, after one round of a detection route whose length (number of hops) is x, the number of nodes that have direction trust is m_d, the number of nodes that have a minimum indirection trust is m_{in}, and the number of nodes that cannot obtain trust is m_{no}; they are

$$m_d = \left(1-(1-\varphi)^x\right)\Big/\varphi,\ m_{in} = (1-P)d,\ m_{no} = (d-m_d)P\ |m<d \quad (17)$$

where

$$\begin{cases} p = 0 & if\ \ \mu < 2\nu\ |\mu = (1-\dfrac{\sqrt{3}}{2\pi})d-2 \\ p = \dfrac{(\mu-\nu)!(\mu-\nu)!}{\mu!(\mu-2\nu)!} & if\ \ \mu \ge 2\nu\ \ |\nu = \left((1-\dfrac{\sqrt{3}}{2\pi})d-2\right)m_d \Big/ d \end{cases}$$

Proof: Considering that the malicious node ratio is φ, the detection route length (number of hops) is x, though during the routing, it may end early due to a black hole. Then, for route length x, the actual average route length is calculated as follows:

The probability of encountering a black hole at the first time is φ, and the probability of not encountering one until the second time is $(1-\varphi)\varphi$; thus, the probability of not encountering one until the i th time is $(1-\varphi)^{i-1}\varphi$.

Thus, the actual average route length is

$$m_d = \varphi + 2(1-\varphi)\varphi + ... + i(1-\varphi)^{i-1}\varphi + ... + x(1-\varphi)^{x-1}\varphi \quad (18)$$

After complex processing, the above equation can be simplified into: $m_d = \left(1-(1-\varphi)^x\right)\Big/\varphi$.

If each node processes one round of detection with length x, then from the average, it is equivalent to for each node to process m_d detection routes to its neighbors; thus, the number of nodes with direction trust is m_d.

For indirection trust, as shown in Fig. 7, node A and node B have a minimum number of common neighbors; then, the indirection trust probability is the minimum it can be calculated as follows:

The number of common nodes of A and B is the number of nodes within the same transmission radius. The area of this region is $2(\pi r^2/3 - \frac{\sqrt{3}}{2}\frac{1}{2}r^2) = \frac{2}{3}\pi r^2 - \frac{\sqrt{3}}{2}r^2$.

The number of nodes in this region is

$$(\frac{2}{3}\pi r^2 - \frac{\sqrt{3}}{2}r^2)\rho = \left(\frac{2}{3}\pi r^2 - \frac{\sqrt{3}}{2}r^2\right)d \Big/ \left(\pi r^2\right).$$

Except for A and B, the number of common neighbors is $\mu = (\frac{2}{3} - \frac{\sqrt{3}}{2\pi})d - 2$.

Node A processed m_d detections, and then the number of detections for the common neighbors is $\nu = \left((\frac{2}{3} - \frac{\sqrt{3}}{2\pi})d - 2\right)m_d \Big/ d$; this also applies to node B. The probabilities of these two sets are completely different, that is,

$$\begin{cases} p = 0 & if\ \ \mu < 2\nu \\ p = \dfrac{c_\mu^\nu . c_{\mu-\nu}^\nu}{c_\mu^\nu . c_\mu^\nu} = \dfrac{(\mu-\nu)!(\mu-\nu)!}{\mu!(\mu-2\nu)!} & if\ \ \mu \ge 2\nu \end{cases}$$

Therefore, the probability that A cannot obtain the indirection trust of B is P, A has d neighbors, among which there are m_d nodes that can obtain direction trust and $m_{in} = (1-P)d$ nodes that can obtain indirection trust, and the number of nodes that cannot obtain trust is $m_{no} = (d-m_d)P\ |m<d$.

■

Theorem 4: If only direction trust is considered, and the number of such nodes is m_d, then the success rate for data packets sent to the sink by nodes that are k hops away is

$$\begin{cases} \chi_d^k = (1-\varphi^{d/3})^{k-1} & if\ \ m_d \ge d \\ \chi_d^k = (1-\varphi^{m_d/3+1})^{k-1} & if\ \ m_d < d \end{cases} \quad (19)$$

Proof: First, calculate the success rate of any of node A's one-hop transmissions. A failed transmission means that node A finds that all of the detected nodes whose hops smaller than itself are black holes; the detected nodes cannot be selected, and A must select from the undetected nodes. If the selected undetected node is a black hole, the transmission fails.

Thus, the failure probability is as follows. There are 3 states for node A, that is, nodes whose hops are larger than, the same as and smaller than A's. For the nodal degree d, the number of nodes whose hops are smaller than A's is $d/3$, and there are $m_d/3$ detections for these smaller nodes, with a total of m_d detections.

If $m_d \ge d$, then all of the neighbors of node A can be detected; then, only if all of the next hop nodes are black nodes can the data transmission fail; the probability of this situation is $\varsigma_1 = \varphi^{d/3}$.

If $m_d < d$, the black node probability for each detection is $\varphi^{m_d/3}$, and the black node probability when choosing the next hop is $\varphi^{m_d/3+1}$; that is, the failure probability is $\varsigma_1 = \varphi^{m_d/3+1}$.

Therefore, for a node at k hops from the sink, if data are sent k hops and the last hop is not a black node, then the success transmission probability for each hop after that is

$$\begin{cases} \chi_d^k = (1-\varphi^{d/3})^{k-1} & if\ \ m_d \ge d \\ \chi_d^k = (1-\varphi^{m_d/3+1})^{k-1} & if\ \ m_d < d \end{cases}$$

■

Inference 4: Considering that the number of nodes with direction trust is m_d and that the number of nodes with indirection trust is m_{in}, the success ratio for nodes at k hops from the sink is

$$\begin{cases} \chi_{di}^k = (1-\varphi^{d/3})^{k-1} & if\ \ |m_d \cup m_{in}| \ge d \\ \chi_{di}^k = (1-\varphi^{(m_d \cup m_{in})/3+1})^{k-1} & if\ \ |m_d \cup m_{in}| < d \end{cases} \quad (20)$$

Proof: Because the indirection trust is within a range of two hops, the black node can be identified with indirection trust, and thus the number of recognizable nodes is the union of direction and indirection nodes, that is, $|m_d \cup m_{in}|$; therefore, Inference 4 can be inferred from Theorem 4.

■

Theorem 5: If only direction trust nodes are considered, and the number of such nodes is m_d, then, for a network whose $R = hr$, the success ratio is

$$\chi_d = \begin{cases} \sum_{k=2}^{h}\left((2k-1)(1-\varphi^{d/3})^{k-1}\right)\Big/h^2 & if\ y \geq d \\ \sum_{k=2}^{h}\left((2k-1)(1-\varphi^{y/3+1})^{k-1}\right)\Big/h^2 & if\ y < d \end{cases} \tag{21}$$

Proof: Theorem 4 gives the success probability χ_d^k for nodes at k hops from the sink because the number of such nodes at k hops is $\rho\left(\pi(kr)^2 - \pi((k-1)r)^2\right) = \pi\rho(2k-1)r^2$.

Then, the number of nodes whose data successfully reaches the sink is $S_k = \pi\rho(2k-1)r^2\chi_d^k$

Because there is no black node within a one-hop range, the total number of packets that successfully arrive at the sink is

$$S_{total} = \sum_{k=2}^{h}\left(\pi\rho(2k-1)r^2\chi_d^k\right) =$$

$$\begin{cases} \pi r^2\rho\sum_{k=2}^{h}\left((2k-1)(1-\varphi^{d/3})^{k-1}\right) & if\ y \geq d \\ \pi r^2\rho\sum_{k=2}^{h}\left((2k-1)(1-\varphi^{m_d/3+1})^{k-1}\right) & if\ m_d < d \end{cases}$$

and the number of packets sent in the entire network is $\pi\rho(hr)^2$. Thus, theorem 5 can be proved.

■

Fig. 8 shows the total data route success ratio with our scheme (only one detection route with a length ϖ =5). As seen, our scheme has a much higher total success ratio than does the shortest routing scheme.

Inference 5: Considering that the number of nodes with direction trust is m_d and that the number of nodes with indirection trust is m_{in} for a network whose $R = hr$, the successful data transmission ratio in our scheme is

$$\chi_{di} = \begin{cases} \sum_{k=2}^{h}\left((2k-1)(1-\varphi^{d/3})^{k-1}\right)\Big/h^2 & if\ \ |m_d \cup m_{in}| \geq d \\ \sum_{k=2}^{h}\left((2k-1)(1-\varphi^{|m_d \cup m_{in}|/3+1})^{k-1}\right)\Big/h^2 & if\ \ |m_d \cup m_{in}| < d \end{cases} \tag{22}$$

Proof: Similar to inference 4, inference 5 can be inferred from theorem 3.

■

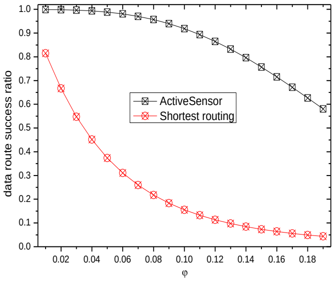


Fig. 8 Total data route success ratio

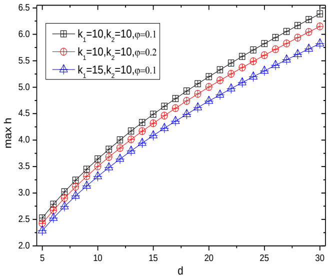


Fig. 9 Required network scale that makes the number of detected nodes larger than the nodal degree without affecting the network lifetime

According to theorems 3-5, if the number of direction detection nodes is larger than the nodal degree, which means that all neighbors are detected and all neighbor trust is obtained, only a scenario in which all neighbors are black nodes can cause the transmission to fail. In fact, if this happens, no scheme can solve this problem because all paths to the sink are blocked by black nodes. Therefore, the situation in which the number of detected nodes equals the nodal degree is optimal. In the following, we analyze whether this ideal situation can be achieved in WSNs.

Theorem 6: For nodal degree d and feedback that is returned hop-by-hop in the detection route, if the network scale meets the following equation, the number of detected nodes can be larger than the nodal degree without affecting the network lifetime, thereby achieving maximum security.

$$h > \sqrt{\frac{3(1-(1-\varphi)^x + 6\varphi)d}{4\varphi}\frac{(1+\kappa_2/\kappa_1)}{(1+\kappa_2)} - \frac{1}{2}} \tag{23}$$

Proof: (1) The energy consumption is the highest in the 1st ring, and the second highest is the 2nd ring. Thus, if the energy can afford the 2nd ring to detect nodes ≥ d, then other rings can ensure that the detected nodes ≥ d. The data load in the 1st ring is $\pi h^2r^2\rho$, and there are $\pi r^2\rho$ nodes in the 1st ring; then, the data load for each node is $\pi h^2r^2\rho\big/\left(\pi r^2\rho\right) = h^2$. Considering that the energy consumption for sending a unit data packet is e_u, the energy consumption in the 1st ring is h^2e_u.

There are $\pi 2^2r^2\rho - \pi r^2\rho = 3\pi r^2\rho$ nodes in the 2nd ring, and the data load is ($\pi h^2r^2\rho - \pi r^2\rho$), so the data load for each node is $\left(\pi h^2r^2\rho - \pi r^2\rho\right)\big/3\pi r^2\rho$ = $\left(h^2-1\right)e_u\big/3$; thus, the remaining energy in the 2nd ring compared with that in the 1st ring is h^2e_u - $\left(h^2-1\right)e_u\big/3$ = $\left(2h^2+1\right)e_u\big/3$. e_p is the energy consumption for a detection packet, so $e_u = (1+\kappa_2)/(1+\kappa_2/\kappa_1)\ e_p$ can be used for detection packets, and the number of detection packets is $\left(\left(2h^2+1\right)\big/3\right)\left((1+\kappa_2)/(1+\kappa_2/\kappa_1)\right)$

If the detection route length is x, then the number of nodes that can be detected is $m_d = \left(1-(1-\varphi)^x\right)\big/\varphi$. If $m_d \geq d$, then there should be at least $n_{\min} = d/m_d$ detection routes.

$$m_d = \left(1-(1-\varphi)^x\right)\Big/\varphi,\ m_{in} = (1-P)d,\ m_{no} = (d-m_d)P\ |m<d$$

For a detection route with length m_d, the number of detection packets needed is

$$1+2+3\ldots+m_d+m_d = \left({m_d}^2+3m_d\right)\Big/2$$

The total number of needed detection packets is

$$n_{\min}\left({m_d}^2+3m_d\right)\Big/2 = \left((1-(1-\varphi)^x+6\varphi)d\right)\Big/(2\varphi)$$

because

$$\left[\left(\left(2h^2+1\right)\Big/3\right)\left((1+\kappa_2)/(1+\kappa_2/\kappa_1)\right)\right] > \left[\left((1-(1-\varphi)^x+6\varphi)d\right)\Big/(2\varphi)\right]$$

$$\Rightarrow h > h > \sqrt{\frac{3(1-(1-\varphi)^x+6\varphi)d}{4\varphi}\frac{(1+\kappa_2/\kappa_1)}{(1+\kappa_2)}-\frac{1}{2}}$$

■

As seen in Fig. 9, if the network scale is only 7 hops with a nodal degree of 30, the residue energy in non-hotspots region can process a sufficient number of detection routes in one round of data collection to detect all neighbors' trust without affecting the network lifetime. This state achieves the best security.

Theorem 7: For nodes that are k hops away from the sink, the success ratio of our scheme when the shortest route is adopted is

$$\beta_k = \chi_d^k\Big/(1-\varphi)^k\ (1-\varphi)^k \qquad (24)$$

Proof: For a black node ratio in the network of φ and for nodes that are k hops away from the sink because nodes are randomly selected, the probability of a black node is the same as the black node ratio in the network, that is, φ, for each hop selection. The last hop is not a black node; thus, with the shortest route scheme, the probability of all non-black nodes being selected after k hops is $(1-\varphi)^{k-1}$, and the ratio of our scheme to the shortest route scheme is

$$\beta_k = \chi_d^k\Big/(1-\varphi)^k\ (1-\varphi)^k$$

■

Theorem 8: In a network whose $R=hr$, the success ratio of our scheme to the shortest route is

$$\beta = \chi_d\cdot h^2\Bigg/\sum_{k=2}^{h}\left((2k-1)(1-\varphi)^{k-1}\right) \qquad (25)$$

Proof: The above theorems have proved that in the shortest route scheme, the success probability of data at k hops to the sink is $(1-\varphi)^{k-1}$. In the network, the number of nodes that are k hops from the sink is

$$\rho\left(\pi(kr)^2-\pi((k-1)r)^2\right) = \pi\rho(2k-1)r^2$$

Thus, in the shortest route scheme, the number of successful data packets at k hops to the sink is

$$S_k = \pi\rho(2k-1)r^2(1-\varphi)^{k-1}$$

There is no black node in the 1st ring; thus, in the entire network, the number of packets to the sink is

$$S_{total} = \sum_{k=2}^{h}\left(\pi\rho(2k-1)r^2(1-\varphi)^{k-1}\right)$$

There are $\pi\rho(hr)^2$ nodes in the network, so the packet success ratio in the entire network is

$$\frac{\pi r^2\rho\sum_{k=2}^{h}\left((2k-1)(1-\varphi)^{k-1}\right)}{\pi\rho(hr)^2} = \frac{\sum_{k=2}^{h}\left((2k-1)(1-\varphi)^{k-1}\right)}{h^2}$$

Therefore, the packet success ratio of our scheme to the shortest route scheme is

$$\beta = \chi_d\cdot h^2\Bigg/\sum_{k=2}^{h}\left((2k-1)(1-\varphi)^{k-1}\right)$$

■

Figs. 10 and 11 show the improved ratio of our scheme to the shortest route scheme. As seen, as the distance from the sink increases, more hops are required for data to be transmitted to the sink, so the success ratio in the shortest route scheme is low; however, our scheme is based on the detected nodal trust, and the success probability is higher because of the selection of high trust nodes. If the black node ratio is higher, it is more improved by our scheme (up to 10 times more), thus confirming the effectiveness of our scheme (see Figs. 10 and 11).

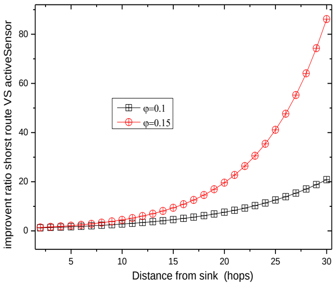


Fig. 10 Improved ratio of our scheme to the shortest route scheme

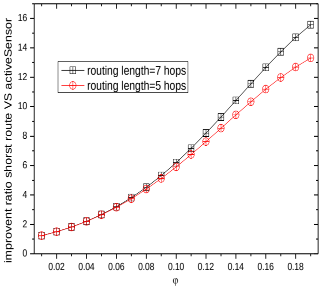


Fig. 11 Total improved ratio of our scheme to the shortest route scheme

G. Analysis of the Energy Efficiency

This section analyzes the energy efficiency performance of our scheme and compares it to other schemes.

Theorem 9: If each node, except for nodes in the 1st ring, processes n_a detection routes with length x, then the energy efficiency is

$$\phi = \sum_{k=2}^{h}\left\{\left(n_d\frac{{m_d}^2+3m_d}{2}e_p+\frac{1-(1-\varsigma_1)^k}{\varsigma_1}e_p\right)\left((2k-1)\right)\right\}\Bigg/\left(h^2e_uh^2\right) \qquad (26)$$

$$|\,n_d = \min\left\{n_a\frac{1-(1-\varphi)^x}{\varphi},d\right\},m_d = \frac{1-(1-\varphi)^x}{\varphi},\varsigma_1 = \varphi^{n_d/3+1}\,|\,\varphi^{d/3}$$

Proof: According to theorem 1, the number of nodes whose direction trust can be obtained in one detection route with length x is $m_d = \left(1-(1-\varphi)^x\right)\big/\varphi\ \left(1-(1-\varphi)^x\right)\big/\varphi$; after n_a detection routes, the number of nodes whose direction trust can be obtained is $n_a = \min\left\{n_a\left(1-(1-\varphi)^x\right)\big/\varphi,d\right\}$. Theorem 4 proved that for a detection route with length m_d, the number of detection packets is $\left({m_d}^2+3m_d\right)\big/2$; thus, for n_a detection routes, the number of detection packets needed is $n_a\left({m_d}^2+3m_d\right)\big/2$. Theorem 2 proved that the number of nodes whose direction trust is available is n_d and that the probability of data failure for the next hop is $\varsigma_1 = \varphi^{n_d/3+1}\,|\,\varphi^{d/3}$. Therefore, for nodes that are k hops from the sink, the number of average data route hops is $\left(1-(1-\varsigma_1)^k\right)\big/\varsigma_1$.

Thus, the energy consumption of a node that is k hops from the sink is

$$n_d \; n_d\left(m_d{}^2+3m_d\right)e_p/2+\left(1-(1-\varsigma_1)^k\right)e_u/\varsigma_1 \; e_p .$$

Because there are $\pi\rho(2k-1)r^2$ nodes that are k hops from the sink, the total energy consumption is

$$\sum_{k=2}^{h}\left\{\left(n_d\,\frac{m_d{}^2+3m_d}{2}e_p+\frac{1-(1-\varsigma_1)^k}{\varsigma_1}e_p\right)\left(\pi\rho(2k-1)r^2\right)\right\}$$

an the highest energy consumption is h^2e_u. Then, the energy efficiency of our scheme is

$$\phi=\sum_{k=2}^{h}\left\{\left(n_d\,\frac{m_d{}^2+3m_d}{2}e_p+\frac{1-(1-\varsigma_1)^k}{\varsigma_1}e_p\right)\left((2k-1)\right)\right\}\Big/\left(h^2e_uh^2\right)$$

■

V. Experimental Results

The experimental platform adopted in this paper is OMNET++ [37]. Unless otherwise noted, the experiments use the following settings. The network radius R =500 m, there are a total of 1000 nodes in the network, among which there are 300 black nodes, nodes are randomly and uniformly deployed, and the sink is at the network's center.

A. *Experimental Results of Node Trust*

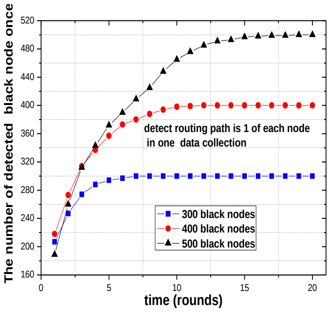


Fig. 12. The number of detected black nodes as the network operates

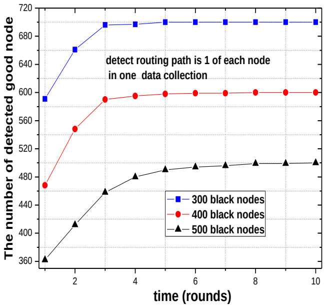


Fig. 13. The number of detected good nodes as the network operates

The experimental scene in Fig. 12 is such that in each data collection round, each node initiates one detection route with a length of 5. As seen, as the network runs, i.e., as more detection routes are performed, the number of black nodes detected grows quickly; when the number of deployed black nodes is 300, 400 and 500, the time needed to detect them all is, respectively, 5, 9 and 12 rounds, which shows that the ActiveTrust scheme can quickly detect malicious nodes within only several detections. Fig. 13 shows the number of detected good nodes as the network runs in the same experimental scene as in Fig. 12; as seen, after only 4 rounds, our ActiveTrust scheme has detected all of the good nodes because in the data routing, it needs only one good downstream node to route the next hop; this indicates that, according to our scheme, the route can be reliable and have a high success probability.

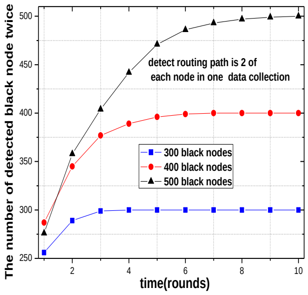


Fig. 14. The number of detected black nodes as the network operates

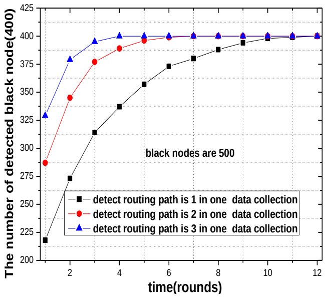


Fig. 15. The number of detected nodes in one data collection round under different numbers of detection routes

Fig. 14 shows the number of black nodes detected in each data collection with twice detecting. As seen, compared with once detecting, the black node detection speed doubles, and all black nodes can be detected in, at most, 7 rounds. The experiment in Fig. 15 further illustrates this problem, which shows that the more detection routes there are in one data collection round, the less time is needed to detect all of the black nodes. This indicates that the black nodes can be more quickly detected as the detection grows, which improves network security. According to inference 2, the residue energy in non-hotspots can afford 7 times (or even more than 10 times) detecting; if all of the residue energy is used to construct detection routes, the system can detect almost all of the black nodes in at most two data collection rounds, which fully verifies the fast recognition ability of our scheme.

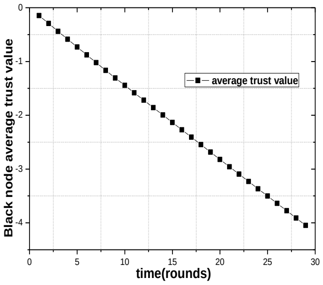


Fig. 16. Average trust of black nodes as the network operates

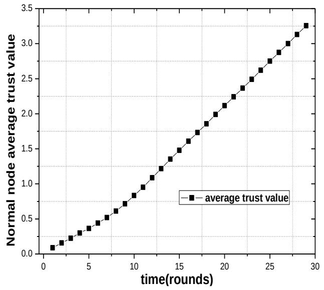


Fig. 17. Average trust of good nodes as the network operates

The experimental scene in Figs. 16 and 17 deploys 1000 nodes in a network with 400 black nodes. In each data collection round, each node creates detection once. As seen, the average trust of black nodes declines as the network operates, whereas that of good nodes increases.

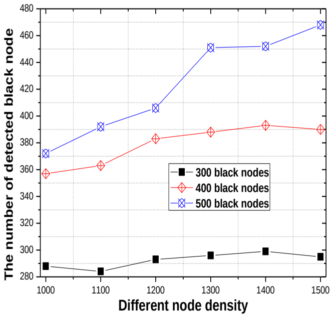


Fig. 18. The number of detected black nodes under different nodal densities

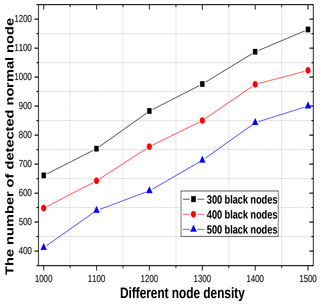


Fig. 19. The number of detected good nodes under different nodal densities

Figs. 18 and 19 show the number of detected black nodes or good nodes after two rounds of data collection when each node detects once in each round for a network of 1000, 1100, 1200, 1300, 1400, and 1500 nodes with 300, 400, and 500 black nodes. As seen from Fig. 18, for a situation with 300 black nodes and a 90% detected black node ratio, the increase in the number of detected black nodes is smaller as the nodal density increases, but if there are 500 black nodes, this increase is more obvious, which shows that our scheme has good performance in networks with greater nodal density. In Fig. 19, the number of detected good nodes grows as the nodal density increases, which shows that in networks with greater nodal density, the success route probability increases, which matches the actual situation.

B. Experimental Results of Energy Consumption

Fig. 20 shows a 3-d map of energy consumption for each node detecting three times in one data collection round in a network with R =400 m and 400 black nodes from a total of 1000 nodes. As seen from Fig. 21, because the detection energy consumption is basically balancing shared, except the detection energy consumption near the sink is very low (to decrease the energy consumption in hotspots), the energy consumption is balanced in other regions; as the detection routes increase, the detection energy consumption increases.

Because the data success route probability is low and most routes are blocked by black nodes in the shortest route scheme, the sink only receives a few data packets. Therefore, in this situation, the energy consumption is more balanced (see Fig. 22). In the ActiveTrust scheme, because the data success route rate is higher, the energy consumption near the sink is higher; although there is detection energy consumption in non-hotspots, the detection energy consumption is low compared with data collection energy consumption, so the energy consumption near the sink is higher than that in other regions. A 3-d map of the energy consumption is shown in Fig. 23, which also indicates that there is sufficient energy remaining for detection.

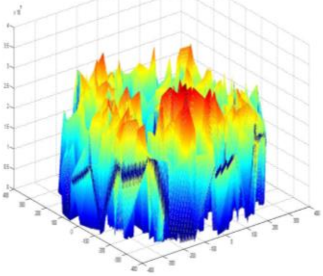


Fig. 20 Energy consumption for each node detecting three times in one data collection round

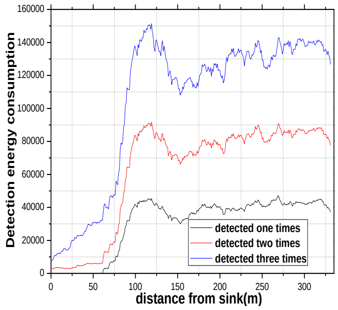


Fig. 21 Detection energy consumption at different distances from the sink

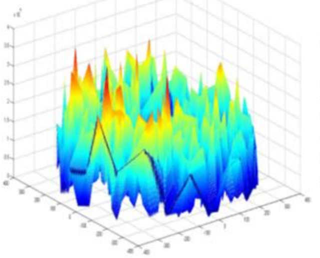


Fig. 22. Energy consumption with the shortest routing scheme

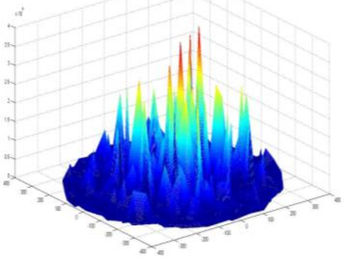


Fig. 23 Detection energy consumption at different distances from the sink

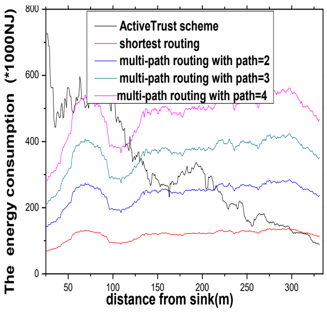


Fig. 24. Energy consumption comparison under different schemes

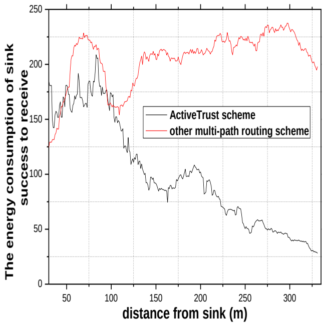


Fig. 25. Energy consumption for unit success under different schemes

Fig. 24 shows the energy consumption at different distances from the sink after one data collection round. As seen, with the shortest routing, the energy consumption is less, as explained previously. With multi-path routing, i.e., one data packet is sent to the sink via different paths to improve the success rate, more packets reach the sink, and the energy consumption is proportional to the number of paths, i.e., the more paths there are, the higher the energy consumption is and the higher the success rate is for data arriving at the sink. Although the success rate increases as the number of paths grows, there are some problems. (1) The success rate is not high; for instance, if the success rate for each path is 20%, then even if 10 paths are created, the success rate does not reach 90%. (2) Even if a certain success rate is achieved, the network lifetime is affected. Therefore, in our scheme, by constructing light active detection routes, malicious nodes can be detected without affecting the network lifetime, which also improves the success rate with good performance.

Fig. 25 shows the ratio of nodal energy consumption to the number of packets that are successfully routed to the sink. This ratio reflects that with the same energy consumption, the number of successful packet in different schemes does, in fact, indicate the network energy efficiency. As seen, our scheme can improve the energy efficiency by more than 2 times compared with that of previous researches which is consistent with theorem 9.

C. Comparison of the Probability of Success Routing

The experimental scene in Fig. 26 is a network with R =400 m and 400 black nodes from a total of 1000 nodes, where each node only detects once. As seen from Fig. 26, as the network runs under our scheme, the probability of successful routing is almost 100% after 7 data collection rounds. For the shortest routing, this probability is not even 15%. With multi-path routing, it is only approximately 60% with 4 paths simultaneously. Moreover, in this black node avoidance scheme, no matter how long the network runs, the probability of successful routing will never increase. The trust-based routing is similar to the TARF scheme [30], in which the next hop is selected based on the trust of the node. Thus, the probability of successful routing will increase with time. However, the scheme does not detect nodes' trust actively, so its probability of successful routing is lower than that of the proposed scheme. Fig. 27 shows the probability of successful routing under different numbers of black nodes. As seen, our scheme is significantly better than multi-path routing. Fig. 28 shows the improvement of our scheme compared with other schemes; as seen, our scheme is better than other schemes. When the network runs a short time, the successful routing probability is improved from 1.5 times to 6 times. Fig. 29 shows the improvement of our scheme compared with other schemes under different numbers of black nodes. As seen, it is improved by more than 3 times compared with the shortest routing and is higher than multi-path routing schemes and trust-based routing.

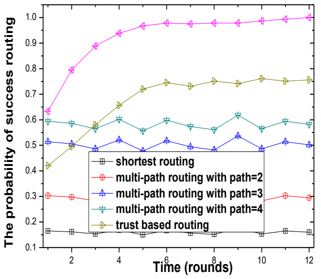


Fig. 26. The probability of successful routing as the network operates

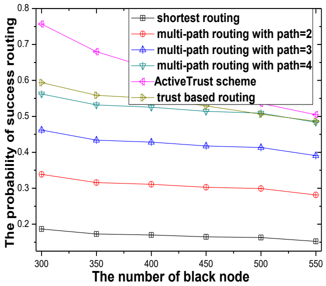


Fig. 27. The probability of successful routing under different numbers of black nodes

Fig. 30 shows the probability of successful routing as the network runs under the ActiveTrust scheme; as seen, even in the situation where there is only one detection in one data collection round, the probability can be almost 100% after several data collection rounds. Fig. 31 shows the probability of successful routing in one data collection round with one, two and three detections. As seen, if the detection routing path is 3, after only 3 rounds, the probability can be almost 100%, which verifies the high probability of successful routing in our scheme.

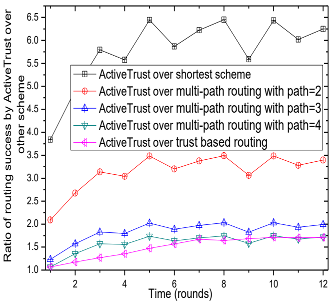


Fig. 28 Ratio of successful routing with different schemes as the network operates

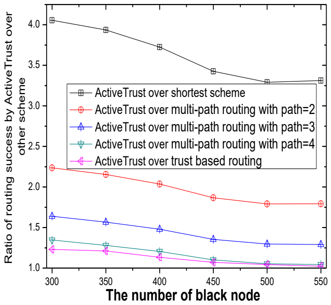


Fig. 29 Ratio of successful routing with different schemes under different numbers of black nodes

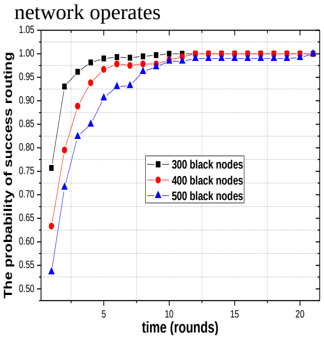


Fig. 30 The probability of successful routing

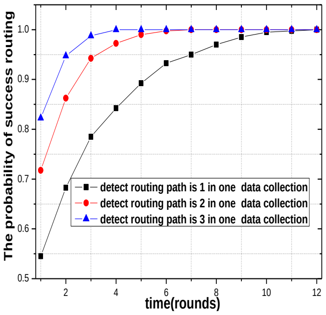


Fig. 31 The probability of successful routing under different numbers of detection routing paths in one data collection round.

Fig. 32 shows the probability of successful routing under different nodal densities. As seen, when the nodal density grows, the nodal degree grows, and the probability of successful routing increases. The reason is that as the nodal density grows, the nodal degree grows, and then there are more detected trustable nodes after detection, that is, there are more nodes for the next hop, and the probability of successful routing thus increases. Fig. 33 shows the probability of successful routing as the nodal transmission radius r grows; as seen, the probability of successful routing is also increased. The reason is that, as r grows, the nodal density grows, which is the same as found in the experiment of Fig. 32.

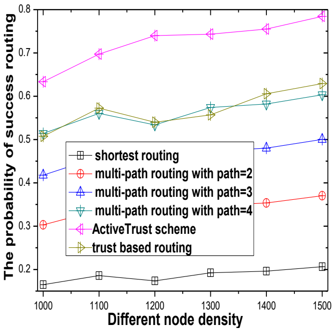


Fig. 32. The probability of successful routing under different nodal densities

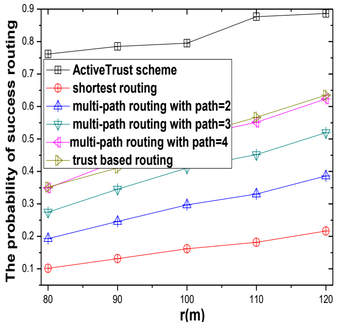


Fig. 33. The probability of successful routing under different nodal transmission radiuses r

Figs. 34 and 35 give the probability of successful routing of the ActiveTrust scheme for different BLAs. In the experiment, the black hole attack refers to the malicious attack in which all data that attempt to pass by are dropped. However, the Denial-of-Service Attack refers to the attack in which data are dropped intermittently [35, 36], thus making it difficult to resist this attack. The select forward attack is one of the most intelligent attacks and can drop data selectively [6]. It can be seen from Figs. 34 and 35 that the ActiveTrust scheme has positive effects on the different impacts of BLAs.

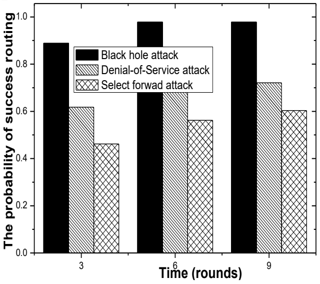


Fig. 34. The probability of successful routing for different BLAs

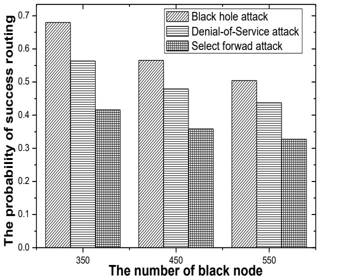


Fig. 35. The probability of successful routing under different numbers of black nodes for different BLAs

VI. Conclusion

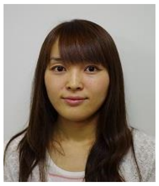
In this paper, we have proposed a novel security and trust routing scheme based on active detection, and it has the following excellent properties: (1) High successful routing probability, security and scalability. The ActiveTrust scheme can quickly detect the nodal trust and then avoid suspicious nodes to quickly achieve a nearly 100% successful routing probability. (2) High energy efficiency. The ActiveTrust scheme fully uses residue energy to construct multiple detection routes. The theoretical analysis and experimental results have shown that our scheme improves the successful routing probability by more than 3 times, up to 10 times in some cases. Further, our scheme improves both the energy efficiency and the network security performance. It has important significance for wireless sensor network security.

REFERENCES

1. Y. Hu, M. Dong, K. Ota, et al. "Mobile Target Detection in Wireless Sensor Networks with Adjustable Sensing Frequency," IEEE System Journal, Doi: 10.1109/JSYST.2014.2308391, 2014.
2. M. Dong, K. Ota, A. Liu, et al. "Joint Optimization of Lifetime and Transport Delay under Reliability Constraint Wireless Sensor Networks," IEEE Transactions on Parallel and Distributed Systems, vol. 27, no. 1, pp. 225-236, 2016.
3. S. He, J. Chen, F. Jiang, et al. "Energy provisioning in wireless rechargeable sensor networks," IEEE transactions on mobile computing, vol. 12, no. 10, pp. 1931-1942, 2013.
4. X. Liu, M. Dong, K. Ota, P. Hung, A. Liu. "Service Pricing Decision in Cyber-Physical Systems: Insights from Game Theory," IEEE Transactions on Services Computing, vol. 9, no. 2, pp. 186-198, 2016.
5. C. Zhu, H. Nicanfar, V. C. M. Leung, et al. "An Authenticated Trust and Reputation Calculation and Management System for Cloud and Sensor Networks Integration," IEEE Transactions on Information Forensics and Security, vol. 10, no. 1, pp. 118-131, 2015.
6. A. Liu, M. Dong, K. Ota, et al. "PHACK：An Efficient Scheme for Selective Forwarding Attack Detecting in WSNs," Sensors, vol. 15, no. 12, pp. 30942-30963, 2015.
7. A. Liu, X. Jin, G. Cui, Z. Chen, "Deployment guidelines for achieving maximum lifetime and avoiding energy holes in sensor network," Information Sciences, vol. 230, pp.197-226, 2013.
8. Z. Zheng, A. Liu, L. Cai, et al. "Energy and Memory Efficient Clone Detection in Wireless Sensor Networks," IEEE Transactions on Mobile Computing.vol. 15, no. 5, pp. 1130-1143, 2016.
9. T. Shu, M. Krunz, S. Liu, "Secure data collection in wireless sensor networks using randomized dispersive routes," IEEE Transactions on Mobile Computing, vol. 9, no. 7, pp. 941-954, 2010.
10. P. Zhou, S. Jiang, A. Irissappane, et al. "Toward Energy-Efficient Trust System Through Watchdog Optimization for WSNs," IEEE Transactions on Information Forensics and Security, vol. 10, no. 3, pp. 613-625, 2015.
11. S. Shen, H. Li, R. Han, et al. "Differential game-based strategies for preventing malware propagation in wireless sensor networks," IEEE Transactions on Information Forensics and Security, vol.9, no. 11, pp. 1962-1973, 2014.
12. O. Souihli, M. Frikha, B. H. Mahmoud, "Load-balancing in MANET shortest-path routing protocols," Ad Hoc Networks, vol. 7, no. 2, pp. 431-442, 2009.
13. J. Long, A. Liu, M. Dong, et al. "An energy-efficient and sink-location privacy enhanced scheme for WSNs through ring based routing," Journal of Parallel and Distributed Computing, vol. 81, pp. 47-65, 2015.
14. S. He, J. Chen, X. Li, et al. "Mobility and intruder prior information improving the barrier coverage of sparse sensor networks," IEEE transactions on mobile computing, vol. 13, no. 6, pp.1268-1282, 2015.
15. S. H. Seo, J. Won, S. Sultana, et al. "Effective key management in dynamic wireless sensor networks," IEEE Transactions on Information Forensics and Security, vol. 10, no. 2, pp. 371-383, 2014.
16. Y. Hu, A. Liu. "An efficient heuristic subtraction deployment strategy to guarantee quality of event detection for WSNs," The Computer Journal, vol. 58, no. 8, pp. 1747-1762, 2015.
17. S. J. Lee, M. Gerla, "Split Multipath Routing with Maximally Disjoint Paths in Ad Hoc Networks," IEEE ICC, pp. 3201-3205, 2011.
18. Y. Zhang, S. He, J. Chen. "Data Gathering Optimization by Dynamic Sensing and Routing in Rechargeable Sensor Networks," IEEE/ACM Transactions on network, doi:10.1109/TNET.2015.2425146, 2015.
19. T. P. Nghiem, T. H. Cho, "A multi-path interleaved hop-by-hop en-route filtering scheme in wireless sensor networks," Computer Communications, vol. 33, no. 10, pp. 1202-1209, 2010.
20. Y. L. Yu, K. Q. Li, W. L. Zhou, P. Li, "Trust mechanisms in wireless sensor networks: Attack analysis and countermeasures," Journal of Network and Computer Applications, vol. 35, no. 3, pp. 867-880, 2012.
21. Q. He, D. Wu, P. K. Sori, "a secure and objective reputation-based incentive scheme for ad hoc networks," IEEE Wireless Communications and Networking Conference, pp. 825–830, 2004.
22. S. Kamvar, M. Schlosser, H. Garcia-Molina, "The eigentrust algorithm for reputation management in P2P networks," in: Proceedings of the 12th International Conference on World Wide Web, pp. 640–651, 2003.
23. H. C. Leligou, P. Trakadas, S. Maniatis, P. Karkazis, T. Zahariadis, "Combining trust with location information for routing in wireless sensor networks," Wireless Communications and Mobile Computing, vol. 12, no. 12, pp. 1091-1103, 2012.
24. J. Wang, Y. H. Liu, Y. Jiao, "Building a trusted route in a mobile ad hoc network considering communication reliability and path length," Journal of Network and Computer Applications, vol. 34, no. 4, pp. 1138-1149, 2011.
25. H. Sun, C. Chen, Y. Hsiao, "An efficient countermeasure to the selective forwarding attack in wireless sensor networks," in Proc. Of IEEE TENCON 2007, pp. 1-4, 2007.
26. Z. Ye, V. Krishnamurthy, S. K. Tripathi, "A Framework for Reliable Routing in Mobile Ad Hoc Networks," Proc. IEEE INFOCOM, pp. 270-280, 2003.
27. W. Lou, Y. Kwon, "H-Spread: A Hybrid Multipath Scheme for Secure and Reliable Data Collection in Wireless Sensor Networks," IEEE Transaction on vehicular technology, vol. 55, no. 4, pp. 1320-1330, 2006.
28. D. R. Stinson. Cryptography, Theory and Practice. CRC Press, 2000
29. Y. Liu, Y. Zhu, L. M. Ni, et al. "A reliability-oriented transmission service in wireless sensor networks," IEEE Transactions on Parallel and Distributed Systems, vol. 22, no. 12, pp. 2100-2107, 2011.
30. G. X. Zhan, W. S. Shi, J. L. Deng, "Design and implementation of TARF: A trust-aware routing framework for WSNs," IEEE Transactions on Dependable and Secure Computing, vol. 9, no. 2, pp. 184-197, 2012.
31. M. Y. Hsieh, Y. M. Huang, H. C. Chao, "Adaptive security design with malicious node detection in cluster-based sensor networks," Computer Communications, vol. 30, no. 1, pp. 2385-2400, 2007.
32. D. He, C. Chen, S. Chan, J. Bu, A. V. Vasilakos, "ReTrust: Attack-resistant and lightweight trust management for medical sensor networks," IEEE Transactions on Information Technology in Biomedicine, vol. 16, no. 4, pp. 623-632, 2012.
33. F. Gómez Mármol, G. Martínez Pérez, "TRIP, a trust and reputation infrastructure-based proposal for vehicular ad hoc networks," Journal of Network and Computer Applications, vol. 35, no. 3, pp. 934-941.2012.
34. G. X. Zhan, W. S. Shi, J. L. Deng J L, "SensorTrust: A resilient trust model for wireless sensing systems," Pervasive and Mobile Computing, vol. 7, no. 4, pp. 509-522, 2012.
35. I. Aad, P. J. Hubaux and W. E. Knightly, "Impact of Denial-of-Service Attacks on Ad-Hoc Networks," IEEE-ACM Transactions on Networking, vol. 16, no. 4, pp. 791- 802, 2008.
36. S. Mandala, k. Jenni, M. A. Ngadi, et al. "Quantifying the severity of black hole attack in wireless mobile ad hoc networks." Security in Computing and Communications. Springer Berlin Heidelberg, 2014: 57-67.
37. OMNet++ Network Simulation Framework, http://www.omnetpp.org/, 2013.

Yuxin Liu Currently she is a student in School of Information Science and Engineering of Central South University, China. Her major research interest are wireless sensor networks, cloud computing.

Mianxiong Dong received B.S., M.S. and Ph.D. in Computer Science and Engineering from The University of Aizu, Japan. He is currently an Associate Professor in the Department of Information and Electronic Engineering at the Muroran Institute of Technology, Japan. His research interests include Wireless Networks, Cloud Computing, and Cyber-physical Systems. Dr. Dong serves as an Editor for IEEE Network, IEEE Communications Surveys and Tutorials, IEEE Wireless Communications Letters, IEEE Cloud Computing, and IEEE Access.



Kaoru Ota was born in Aizu Wakamatsu, Japan. She received M.S. degree in Computer Science from Oklahoma State University, USA in 2008, B.S. and Ph.D. degrees in Computer Science and

Engineering from The University of Aizu, Japan in 2006, 2012, respectively. She is currently an Assistant Professor with Department of Information and Electronic Engineering, Muroran Institute of Technology, Japan. She serves a Guest Editor of IEEE Wireless Communications, Her research interests include wireless sensor networks, vehicular ad hoc networks, and ubiquitous computing.

Anfeng Liu is a Professor of School of Information Science and Engineering of Central South University, China. He received the M.Sc. and Ph.D degrees from Central South University, China, 2002 and 2005, both in computer science. His major research interest is wireless sensor network, Crowd sensing network.