IEICE TRANS. INF. & SYST., VOL.E99–D, NO.8 AUGUST 2016

PAPER Special Section on Security, Privacy and Anonymity of Internet of Things

MDMA: A Multi-Data and Multi-ACK Verified Selective Forwarding Attack Detection Scheme in WSNs

Anfeng LIU[†], **Xiao LIU**[†], ***Nonmembers***, **He LI**[††], ***Student Member***, ***and*** **Jun LONG**[†a)], ***Nonmember***

SUMMARY In this paper, a multi-data and multi-ACK verified selective forwarding attacks (SFAs) detection scheme is proposed for containing SFAs. In our scheme, each node (in addition to the nodes in the hotspots area) generates multiple acknowledgement (ACK) message for each received packet to confirm the normal packet transmission. In multiple ACK message, one ACK is returned along the data forwarding path, other ACKs are returned along different routing paths, and thus malicious nodes can be located accurately. At the same time, source node send multiple data routing, one is primary data routing, the others are backup data routing. Primary data is routed to sink directly, but backup data is routed to nodes far from sink, and then waits for the returned ACK of sink when primary data is routed to sink. If a node doesn't receive the ACK, the backup data is routed to sink, thus the success rate of data transmission and lifetime can be improved. For this case, the MDMA scheme has better potential to detect abnormal packet loss and identify suspect nodes as well as resilience against attack. Theoretical analysis and experiments show that MDMA scheme has better ability for ensuring success rate of data transmission, detecting SFA and identifying malicious nodes.

key words: *Wireless Sensor Networks, selective forwarding attack, multi-data and multi-ACK, identify malicious node*

1. Introduction

Wireless Sensor Network (WSNs) are being emerged as one of prevailing technology due to its wide range of applications in military and civilian domains such as battlefield surveillance, medical monitoring, biological detection, etc [1]–[5]. Due to their operating nature, they are often unattended, hence prone to different kinds of attacks [3]–[8]. Selective forwarding attack is one of such attacks [3]–[5], it intelligently drops some packets to damage network, and SFA is simple to implement but difficult to detect attack behavior [5]. A multi-data and multi-ACK (MDMA) verified selective forwarding attacks detection scheme is proposed to contain SFAs, the innovation points in this paper are mainly:

(1) The MDMA scheme proposed in this paper has higher security ability, and better ability for detecting and identify suspect nodes. In MDMA scheme, each intermediate node along a forwarding path is responsible for generating multi acknowledgements (ACK) to source node for each packet received. In multiple ACKs, an ACK is routed to source node along the forwarding path (this ACK identified as ACK_0), but other ACKs are routed to source node along different routing paths (those ACKs identified as ACK_1, ACK_2, ... ACK_k, using ACK_{other} to express other ACKs except ACK_0). The ACK_0 is used to detect malicious node in forwarding path, even if malicious node drop the ACK_0, the ACK_{other} are also routed to source node along different paths. So the probability which only an ACK is routed to source node is higher, if the malicious node drops ACK_0, it not only doesn't damage network, but also will expose themselves, so as to identify malicious node. Thus this scheme has higher ability for detecting and identifying suspect nodes.

(2) Second, in MDMA scheme, the different from previous schemes are: it not only can effectively detect selective forwarding attack, but also has the function of recovering failed routing fast. Multiple data routing are produced at the same time, the data in other path is routed to a certain location, which is used to detect whether primary data is routed to sink successfully, if it isn't, the routing continues to route to sink, so the success rate of routing is greatly improved, and greatly reduced the danger of SFA attack.

(3) MDMA scheme has higher network lifetime. Most proposed schemes which have better safety performance also need to consume much energy. But MDMA scheme not only have better safety performance, but also have higher lifetime. The reason is: in the process of data collection, the nodes in the range of one hop from sink need to receive and forward data packets from nodes far away from sink, the energy consumption is much higher in this area than the energy consumption in other areas, it called hotspots, when the energy consumption of the nodes in hotspots area is used up, this leads to sink can't receive data packet from nodes far away from sink, which make the network died in advance [1], [3]. In MDMA scheme, it sends multiple data and ACK flows to make full use of the energy of area far away from sink, the energy consumption in hotspots area is the same with previous schemes. Thus the network lifetime isn't reduced, but detecting accuracy and effectiveness of the proposed protocol in this scheme has improved a lot.

(4) We provide theory analysis and extensive simulation results for MDMA scheme. The experimental results confirm the validity of this scheme. The detecting probability and the accuracy of identifying suspect nodes in this scheme are better than previous studies.

Manuscript received November 19, 2015.
Manuscript revised April 9, 2016.
Manuscript publicized May 31, 2016.

[†]The authors are with School of Information Science and Engineering, Central South University, ChangSha, 410085, China.
[††]The author is with Computer Science and Engineering, Huazhong University of Science and Technology, Wuhan, 430074, China.
a) E-mail: longjuncs@sina.com

DOI: 10.1587/transinf.2015INP0005

Copyright © 2016 The Institute of Electronics, Information and Communication Engineers

2. Related Works

The selective forwarding attack caused great harm to the network, thus there are some researches about detecting and resisting selective forwarding attack.

(1) Multi-path routing scheme. The main method of multi-path routing scheme is that the same data packet are routed to sink through multiple routing paths. The routing failed only when all these routing paths are attacked, the probability which those routing paths are attacked at the same time is lower, so the ability for resisting attacks is improved. Multi-path routing scheme is not only for SFA, but it has better performance for SFA. Karlof et al. [9] first discuss the selective forwarding attack and also suggest that multi-path routing can be used to counter these types of attacks. The advantage of this scheme is that the implementation is simple, it also has a broad ability of resisting security attacks. Its defect are: there is no corresponding attack detection mechanism, the location of the attacked nodes can't be identified when data packets are dropped. The more serious is that its energy consumption is much times than that of the single routing scheme, thus seriously affects network lifetime.

(2) Subsequent schemes can detect whether there is SFA, but the ability of identifying malicious nodes is weak. In this paper, these schemes are no longer just for routing data, but has the ability for detecting attacks, if the data routing is attacked, it will send warning information to the system, which warning the following routing don't enter the attack area any more. This kind of scheme has weak ability of identifying malicious nodes. Hung-Min Sun et al. [10] have proposed a multi dataflow topologies (MDT) method to countermeasure the selective forwarding attack, in this way. In multiple MDT, as long as a data flow don't be attacked, the data packet can be routed to sink successfully. The failure of dataflow shows that there are malicious nodes. This kind of scheme has ability for identifying malicious nodes, but the ability for identifying malicious nodes is weak, the biggest problem is the energy consumption is also several times than that of single routing scheme, and serious damage lifetime [11].

(3) The proposed scheme has better ability for ensuring data routing security, detecting and identifying malicious nodes. In terms of detection SFA, Xiao, Yu et al. [5] have proposed a CHEMAS (checkpoint-based multi-hop acknowledgement scheme) to against SFA. The main point of CHEMAS scheme is that it chooses a certain number of nodes as checkpoint nodes in the routing path from source node S to sink. Checkpoint node will return ACK to upstream node of the routing path when it receives data packet, an ACK packet contains survive time of ACK packet, that is, time to live (TTL), and the number of TTL minus 1 when the ACK packet pass a checkpoint node, if the number of TTL is 0, then the data packet is discarded. After nodes forward data, it just wait for the arrival of ACK packet, if nodes didn't receive the expected number of ACK packet, it will send alert message to source node.

3. The System Model

3.1 The Network Model

(1) We consider a wireless sensor network consisting of n sensor nodes that are uniformly and randomly scattered in a circle network, the network radius is R, with the density of nodes ρ, and nodes do not move after being deployed [12]. On detecting an event, a sensor node will generate a data packet and the data packet need to be sent to sink node [2], [13]. The shortest routing approach is considered to employ in this paper [2], [13], [14].

(2) The attacker is considered has a strong intelligence [5]. It obtains the legal status of these nodes through compromising a small portion of sensor nodes, then drop some data packets, ACK message or alter message at a certain probability, the aim is to try not to expose themselves, and to make the greatest harm to the network. At the same time, the attackers can also collude to launch attacks.

(3) Message authentication code is adopted in MDMA scheme which provides assurance to the recipient of the message came from the expected sender and has not been altered in transit [6]. Therefore, in this paper, if there is no special instructions, all packets or message adopt message authentication code technology.

3.2 Energy Consumption Model and Related Definitions

Adopting the typical energy consumption model [2], [12], [13], see energy consumption for sending data in Eq. (1) and for receiving data in Eq. (2).

$$\begin{cases} E_t = lE_{elec} + l\varepsilon_{fs}d^2 & if\ d < d_0 \\ E_t = lE_{elec} + l\varepsilon_{amp}d^4 & if\ d > d_0 \end{cases} \tag{1}$$

$$E_r(l) = lE_{elec} \tag{2}$$

E_{elec} in the formula means the energy consumption of transmitting circuit. If the transmitting distance is less than threshold d_0, the consumption of power amplification adopts the free space model. If the transmitting distance is more than threshold d_0, adopts the multipath attenuation model. ε_{fs} and ε_{amp} are the energy required to amplify power respectively in the two models. l denotes the number of bits of data. In this paper, the parameter of specific configuration

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

above references Refs. [2], [12], [13], shown as Table 1.

4. The Design of the Protocol

4.1 Research Motivation

The key problems of designing selective forwarding attack detection scheme are: (1) the capability of resisting security attacks. In theory, once SFA is attacked, data packet will be dropped by attacker, which resulting in routing failure. In order to protect data packet to reach sink successfully when date packet is attacked, this paper adopts the redundant multi-path routing mechanism, if multi-path routing mechanism creates β routing path, the network lifetime in multi-path routing mechanism is $1/\beta$ times than that of single routing scheme, and the cost is unbearable of wireless sensor network, and thus limits the use of redundant routing. (2) The ability for identifying malicious nodes. In order to know arrived location of data transmission, the scheme usually uses the method which nodes return ACK to source node, at the same time, it has the ability of identifying malicious nodes. But, the challenge issue of this scheme are: the more the returned ACK are, the better ability for identifying malicious nodes is, but this can affect the network lifetime. However, the previous schemes select a certain of nodes to return ACK [5]. The another problem is: all ACKs are returned to source node along data forwarding path in the previous schemes, but if there are malicious nodes in the routing path, it can also drop ACK packet or alter message to make the scheme failure.

Therefore, the main improvements in MDMA scheme are: (1) for data packet, it uses multi-path routing scheme. Thus it can ensure higher ability for resisting attack when under attack; (2) The node not only return an ACK along data forwarding path when receives data packet, but also return multiple ACKs along other routing paths, so it can improve the ability for resisting attack. It shows that the MDMA scheme have better performance in security and identifying malicious nodes. Moreover, MDMA scheme make full use of the residual energy in non-hotspots area, this paper gives the number of multiple data routing through the theoretical calculation, as well as the calculation method of routing length of backup routing, this can ensure the network lifetime equal to the single data path and single ACK path scheme under the situation of creating multiple data path and ACK path. Because MDMA scheme make full use of more than 90% of the surplus energy to resist SFA, it has good performance.

4.2 Overview of the Proposed Scheme

The overview of the MDMA scheme is shown as Fig. 1, the main components are as follows:

(1) Data route path. The main goal of MDMA scheme is that the data is routed to sink successfully. In MDMA scheme, the data routing is not only a routing path, but have multiple routes at the same time, including a route known as primary data route, other routes known as backup data routes. Primary data is routed using the shortest routing, because the network exists selective forwarding attack, primary data packet may be dropped. Backup data routes are established to resist data attack in MDMA scheme. Backup route is similar to multi-path routing scheme, it also set up multiple routing, but the difference from previous multi-path routing scheme is: Backup route only builds routing in non-hotspots area, it does not establish a direct routing to sink, only if the primary route can't reach sink successfully, it will continue to route to sink. It can achieve two goals: (1) if the primary route fails, backup route can continue to be routed to sink, so as to resist attack, it has very good security. (2) The higher network lifetime. In multiple routing, all routes are routed to sink, and the network lifetime is only $1/\beta$ than that of single routing (if the network don't be attacked), and thus reduce the network lifetime. And MDMA broke this limit, the ability for resisting attack is almost the same as the multi-path routing scheme, but the network lifetime is almost the same as the single routing scheme. The reason is: although MDMA creates more than one backup route, those backup routes only are routed to sink when primary route isn't routed to sink successfully. Otherwise those backup routes is routed to the node outside hotspots (the node is called anchor node). Those backup routes are not pass through hotspots area. So the appropriate design of backup route can make full use of the residual energy in non-hotspots. As a result, it have good sense.

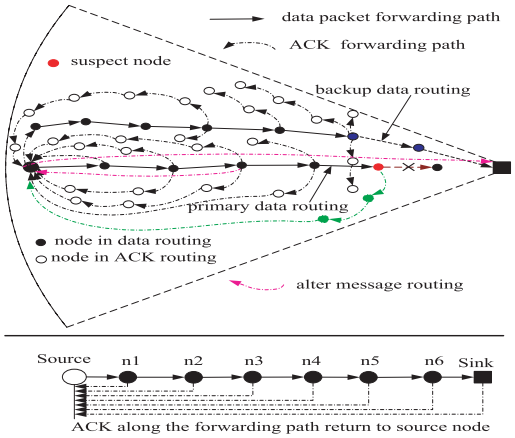


Fig. 1 Illustrate of the MDMA scheme

(2) ACK route path. It can be seen through analysis that: although the energy of the wireless sensor network (WSN) is very limited, the residual energy of non-hotspots is more than 90%, which can not only realize the data route, but also realize each intermediate nodes along a forwarding path is responsible for generating multi acknowledgements (ACK) to source node for each packet received. Among them, primary ACK_0 along the forwarding path return to source node, ACK_{other} is independent routed to source node along different routing paths. Attack behavior of malicious node can be identified by cross recognition of ACKs. Thus MDMA scheme has higher ability for detecting and identi-

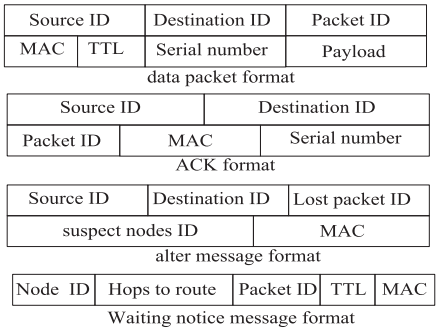


Fig. 2 Packets format

fying suspect nodes.

4.3 The Format of Data Packets and Message

In MDMA scheme, there are four kinds of different packets flow, its format such as Fig. 2. (1) Data flow. It is produced by source node, which contains seven domains: (a) source ID is the identification (ID) of source node which produces data packet. (b) Destination ID refers to the destination node ID of data packet, it is sink. (c) Packet ID, data packet ID, each data packet has a unique ID, it is used to identify data packet. (d) MAC refers to Message authentication code (MAC). (e) TTL, namely time to live, it is the time which the node holds data packet. In MDMA, primary data is directly routed to sink, the intermediate node passed by data packet don't need to save data packet, and for the last node in backup data routing, i.e. anchor node, it drops data when it receives the returned message which the primary data is routed to sink successfully, or it re-launch the routing when it doesn't receive the returned message. Therefore, TTL refers to the holding time that anchor node holds data packet in backup data routing. The value of TTL minus 1 in a while, the time is requirement time of which the data is routed for one hop, if the value of TTL is zero, anchor node drop data. (f) Serial number. It refers to the number of data route path. It shows primary data route when Serial number is 0, and shows backup data route if the value is bigger than 1. (g) Payload. Refers to the content of data packet.

(2) ACK format. In ACK format, Source ID refers to ID of node which generate ACK. Destination ID refers to destination node ID of ACK, generally refers to ID of source node. Packet ID said ID of the received data packet, as a node returns multiple ACKs when it receives data packet, it represents that a returned ACK is ACK of this data packet. MAC is Message authentication code (MAC). Serial number is the number of ACK route path. It shows that the returned ACK is routed along the data forwarding path when Serial number is 0, and shows that the routing path to confirm ACK which is returned along other routing path when Serial number is bigger than 1.

(3) Alter message format. Alter message format have two kind of situations: (a) after a node in the network identifies some suspect nodes, in order to tell source node that there may be malicious node, alter message is routed to source node by this node; (b) source node report alter message of suspect nodes to sink after those alter message are integrated. Sink can take measures to clear and verify malicious nodes. Alter message added two fields: (a) Lost packet ID is refers to ID of the lost packet; (b) Suspect nodes ID is refers to ID of the suspicious node.

(4) Waiting notice message format. Waiting notice message is used in waiting notice routing which is initiated by anchor node in backup routing. In order to obtain ACK message about whether data routing is routed to sink successfully, the scheme uses the method that anchor node initiate route vertical to data routing, so ensure that waiting notice routing can meet the returned ACK message of data routing from sink, which can obtain message that data routing is successfully routed to sink. It consists of five fields, (a) node ID is ID of sponsor node of notice routing, which is ID of anchor node. (b) Hops to routing indicates routing length of notice routing, the value minus 1 when the data is routed for one hop, it stops routing when the value is 0. Other fields are the same as earlier.

4.4 Routing in MDMA

In MDMA scheme, the main rout has 3 kinds, one is data routing. The main data routing has: (a) primary data routing, using shortest routing method to send data to sink. (b) Backup data routing. Considering produce g data routing. The possible routing can be numbered by source node as $\{1, 2, \ldots, g\}$, the data in the data path numbered i is routed to anchor node which the distance from sink is q_i hop. The formation process of data routing in path numbered i is as follows: source node select the node which is in the left and the same hop count from sink as the first node to route data according to the rules of the left hand, for the data path numbered i, the data is routed for w_i hop count toward left direction, then launch the shortest routing to sink, the data in the path numbered i is routed to anchor node. Then, the anchor node initiates vertical routing in the direction of data routing, it called waiting notice routing. The routed message format is shown as the waiting notice message format of Fig. 2. The routing mechanism of waiting notice routing is given in Fig. 3. After the anchor node initiates the routing to form notice path, each node in the notice path are waiting to see if there is ACK returned from sink which is used to confirm whether sink receives data packet and whether include packet ID in waiting notice message. If it receives this ACK, the node will report the message which the data is routed to sink successfully to anchor node along reverse notice routing path. Otherwise, after the anchor node waiting for a certain of time, the data of anchor node continue to route to sink. It still comply with the mechanism that each hop returns multiple ACKs message in the process of routing. Its route algorithm is shown in algorithm 1.

(2) ACK flow routing. The formation of ACK flow routing is shown as Fig. 2, it includes two categories: the routing between ACK_0 and ACK_{other}. (a) The routing of

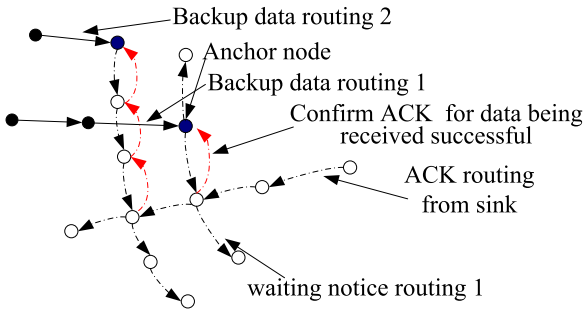


Fig. 3 Intersection of ACK and waiting notice routing

Algorithm 1: The data routing mechanism

Initialize : Each node get its hop to sink using hop diffusion protocol [7].
Node a generate a data packet for sensing event ;

Primary data routing

$\{a.ID, \sin k.ID,$ data.ID, MAC, 0, 0,content$\} \rightarrow \Omega_0$;
$n_0 = a$;
While n_0!=sink
  B=the least hop to sink neighbour node of n_0
  n_0 send Ω_0 to nodes B;
  n_0=B;
End while;

Backup data route path i

$\{a.ID, \sin k.ID,$ data.ID, MAC, ?, i,content$\} \rightarrow \Omega_i$
$n_0 = a$; $x = i\partial$; j =0;
While $j < x$ Do //this is same hop routing
  B= the farthest neighbour node of n_0 with same hop to sink;
  n_0 send Ω_i to nodes B;
  n_0=B; j++;
End while
j =0
While $j < q_i$ Do // Shortest routing to anchor node
  B= the neighbour node of n_0 with the least hop to sink;
  n_0 send Ω_i to nodes B;
  n_0=B; j ++;
End while

Waiting notice routing

Anchor node n_i of backup data path i generate a notice message $\wp$;
Let j =0; υ is the hops to route of $\wp$; n_0=n_i
While $j < \upsilon$ Do
  B= the left (right) neighbour node of n_0 with the same hop to sink;
  n_0 send $\wp$ to nodes B;
  n_0=B; j++;
End while
Once the node receive a ACK from sink Do
  If the packet ID in ACK is same as the $\wp$'s;
    generate a notice message to anchor node n_i;
    anchor node n_i unload notice message $\wp$
  End if

ACK_0 is simple, it is routed to source node along data forwarding path. (b) The formation process of ACK_{other} is as follows: the produced ACK of nodes random walk ϑ hop in the direction away from nodes, then those ACKs are routed to source node using the shortest routing scheme. Please see algorithm 2.

Algorithm 2: The ACK routing mechanism

Initialize : If node a receive a data packets;
node a produce ν ACK message $\xi_i \mid i \in \{0..\nu-1\}$.
$n_0 = a$,
B= the upstream neighbour node data forwarding path of n_0;

ACK_0 routing

While B is not source node Do
  Send ACK_0 to node B;
  B=B's upstream node of forwarding path;
End Do
Send ACK_0 to source node;

ACK_{other} routing

For each ACK_{other} Do
node a random walk hop $\vartheta_i \mid i \in \{0..\nu-1\}$
  B is the neighbour node of node a
  While j< ϑ_i Do //The initial value of j is 1
    Send ACK_i to node B
    B= B's neighbour node which is farthest from node a
    j = j +1;
  End while
  While B is not source node Do
    Send ACK_i to B
    B= B's neighbour node which nearest to source node
  End while
End for

(3) Alter message routing. It's routing is relatively simple, alter message is routed to sink along reliable routing path which the data is routed to sink successfully before. So it has higher success rate of data transmission along this path.

The ACK routing algorithm 2 is shown above.

4.5 Identification of Suspect Nodes

For the sake of simplicity, the network is considered under ideal radio conditions, so the packet loss must be the result of malicious dropping. In MDMA scheme, source node collect the information of all nodes in routing paths, if the source node detect that there is suspect node, than send alter message to sink. Non-source node send alter message to source node when it detects abnormal, there are following situations that non-source node produce alter message: (1) node a, it is in the data forwarding path and the distance from sink is i hop. If the number of received ACK is less than $i-1$, it shows that the routing path from node a to sink exists anomalies. The reason is that the data may be dropped, or the returned ACK may be dropped by malicious node, Thus report anomalies to source node; (2) Anchor node initiate data routing and send alter message to source node if it doesn't receive the message which data reach to sink successfully. After the source node collects ACK message and alter message, the source node identify suspect nodes and forms alter message, then those are sent to sink. For example, the following circumstances may be suspect node: (a) the next-hop node and the last node for returning ACK node. (b) The nodes which the received ACK

don't arrive the expected number.

5. Parameter Optimization and Performance Analysis

5.1 Parameter Optimization

Considering network radius $R = \hbar r$, the hop count between source node S and sink is h, parameter optimization in MDMA scheme has the following theorem.

Theorem 1: In MDMA scheme, considering the nodes in i th ring create g_i data paths, the nodes return κ ACKs to the source node when the nodes receive a data packet. The energy consumption for forwarding data packets and ACK packets from the nodes in other rings is:

$$E_i = \lambda e_u \left(\frac{g_i(m^2 - i^2)}{(2i-1)} + g_i + \frac{\sum_{j=1}^{i} \{\kappa g_j (m^2 - i^2)(1+\varepsilon)\}}{(2i-1)k} \right) \tag{3}$$

Proof: Considering the length of data packet is l bits, the length of ACK is δ bit, set $k = l/\delta$. The energy consumption for forwarding a data packet is e_u. The production rate of event is λ, so the amount of data loaded by the nodes is as follows:

The routing path of the nodes in $> i$ th rings must pass the nodes in i th ring, the number of total nodes in $> i$ th rings are: $\pi\hbar^2 r^2\rho - \pi i^2 r^2\rho$, the number of data packet need to be loaded by the nodes in i th ring are: $\pi r^2\rho\lambda(\hbar^2 - i^2)$. The number of creating data path is g_i when each data packet arrive to the nodes in i th ring, the number of data packet need to be forwarded by the nodes in i th ring are: $g_i\pi r^2\rho\lambda(\hbar^2 - i^2)$, the number of total nodes in i th ring are:

$$\pi i^2 r^2\rho - \pi(i-1)^2 r^2\rho = \pi(2i-1)r^2\rho$$

Therefore, the energy consumption of the nodes in i th ring for forwarding those data from other rings is:

$$\frac{g_i\pi r^2\rho\lambda(\hbar^2 - i^2)e_u}{\pi(2i-1)r^2\rho} = \frac{g_i(\hbar^2 - i^2)}{(2i-1)}\lambda e_u$$

In MDMA scheme, each node returns κ ACK packets, all produced ACK packets by the nodes in $\geq i$ th rings are forwarded by the nodes in i th ring, the number of data packet loaded by those nodes are: $\pi r^2\rho\lambda(\hbar^2 - i^2)$, each node produces g_j data routs in $j \mid j \leq i$ th ring. The returned path of the produced ACK by each routing path are: κ, the produced ACK by each routing path need to be forwarded by the nodes in i th ring, the ACK packet loaded by those nodes is:

$$\sum_{j=1}^{i} \{\kappa g_j \pi r^2\rho\lambda(\hbar^2 - i^2)(1+\varepsilon)\} \tag{4}$$

Therefore, the energy consumption of the nodes in i th ring for forwarding data packet and ACK packet from other rings is:

$$E_i = \frac{g_i(\hbar^2 - i^2)}{(2i-1)}\lambda e_u + g_i\lambda e_u + \frac{\sum_{j=1}^{i} \{\kappa g_j \pi r^2\rho\lambda(\hbar^2 - i^2)(1+\varepsilon)\} e_u}{\pi(2i-1)r^2\rho k}$$ ■

Theorem 2: In MDMA scheme, the created routing path is g_i by the nodes in i th ring, it meets the following formula, the network lifetime in MDMA scheme is the same as the previous schemes.

$$\left(\frac{g_i(\hbar^2 - i^2)}{(2i-1)} + g_i + \frac{\sum_{j=1}^{i} \{\kappa g_j (\hbar^2 - i^2)(1+\varepsilon)\}}{(2i-1)k} \right) \leq \left(\hbar^2 + \frac{\kappa(\hbar^2 - 1)(1+\varepsilon)}{k} \right) \tag{5}$$

Proof: In MDMA scheme, the way which the nodes in the hotspots area send data packet is the same with previous schemes, so the energy consumption in hotspots is the same with previous schemes. Therefore, as long as the energy consumption of the nodes in $i > 1$ rings is not more than the energy consumption of the nodes in the first ring, the network lifetime is the same with previous schemes. When $i = 1$, the energy consumption is:

$$E_1 = \lambda e_u \left(\hbar^2 + \frac{\kappa(\hbar^2 - 1)(1+\varepsilon)}{k} \right)$$

According to theorem 1, the energy consumption of the nodes in i th ring is E_i, so as long as $E_i < E_1$, the network lifetime is the same with the previous schemes. ■

5.2 Analysis Probability for against Attacks

Considering there are n sensor nodes in whole network, m of which are malicious nodes. We suppose that the hops count between source node and sink is $\hbar$, so, there are $c = \hbar m/n$ malicious nodes in the forwarding path to sink.

Theorem 3: In MDMA scheme, the ratio of the probability which ACK is routed to the source node successfully in MSMA scheme and the previous schemes is:

$$\phi_{ACK} = (1 - (1-\tau)^{g \cdot m_1}) \,/\, (1 - (1-\tau)^{m_1}) \tag{6}$$

Proof: Considering the length of the returned path of ACK is h count hop, so, there are $m_1 = hm/n$ malicious nodes in that path. Considering malicious nodes drop ACK at a random probability τ. So, if those m_1 malicious nodes don't drop ACK, ACK can reach source node. So in the scheme which only a ACK are returned to source node, the probability which the ACK can be routed to source node successfully is: $p^1_{ACK} = 1-(1-\tau)^{m_1} \mid m_1 = hm/n$. But in MDMA scheme, the node return g ACK to source node, the probability which

an ACK reach to source node is: $p^2_{ACK} = 1-(1-\tau)^{g.m_1}$, so it can be proved. ■

Theorem 4: In MDMA scheme, the ratio of the probability which data packet reach sink successfully in MDMA scheme and previous schemes is:

$$\phi_{ACK} = (1-(1-\tau)^{g.m_1}) / (1-(1-\tau)^{m_1}) \quad (7)$$

Proof: Considering the length of the routing path from the source node to sink is $\hbar$ hop count. In previous scheme, the probability for dropping data: $p^1_{data} = 1-(1-\tau)^{m_1} \mid m_1 = \hbar m/n$. But in MDMA scheme, there are g data routing path, the probability which there is a data to reach to sink: $p^2_{data} = 1-(1-\tau)^{g.m_1}$, it is proved in this way. ■

6. Experiment Results

OMNET++ is used for experimental verification [15]. The contrast graph of the energy consumption in MDMA scheme and ACK based scheme is respectively given in Fig. 4 and Fig. 5. It is more balanced in MDMA scheme than that of ACK based scheme. The energy consumption and network lifetime are shown respectively in Fig. 6 and Fig. 7. From Fig. 6, the maximum energy consumption in MDMA scheme is lower than that of ACK based scheme, and the nodes in non-hotspots area consume more energy in MDMA scheme than that of ACK based scheme. It can been seen from Fig. 7 that the network lifetime in MDMA scheme is higher than that of ACK based scheme, the main reason is that though the data packet is routed to sink along different paths in MDMA scheme, the data in most of paths is routed to anchor node which is in the non-hotspots area, only primary path is directly transmitted to sink, if the data in primary path isn't transmitted to sink successfully, the data in other path can be routed to sink. It isn't damage the network lifetime. It has better performance.

The energy consumption in MDMA scheme with different models and the number of data paths in different ring are given in Fig. 8 and Fig. 9. It can be seen that if the nodes which is far from sink have much data paths, this area can consume much more energy than the nodes in hotpots area, the reason is: the node has more data paths, the more data can be transmitted to those nodes, and the more ACK can be returned by those nodes, the much energy can be consumed for transmitting data and ACK. If the nodes have too much data paths, its energy consumption in this area is more

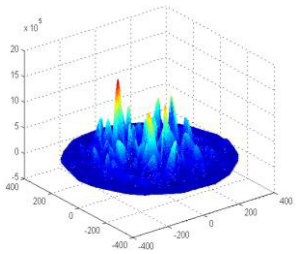


Fig. 4 The energy consumption in MDMA scheme

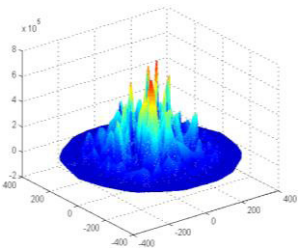


Fig. 5 The energy consumption in ACK based scheme

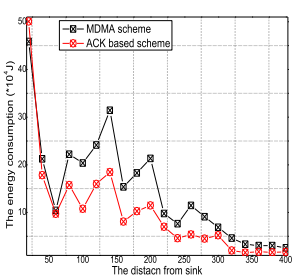


Fig. 6 The energy consumption in different schemes

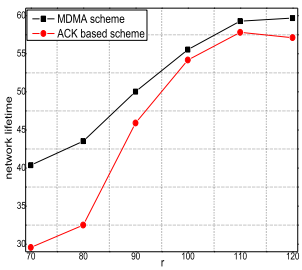


Fig. 7 The network lifetime in different schemes

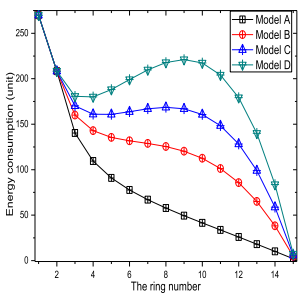


Fig. 8 The energy consumption in MDMA scheme with different models

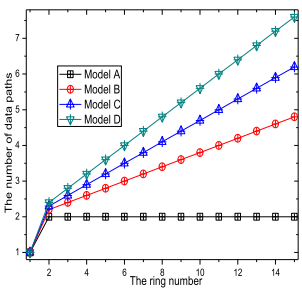


Fig. 9 The number of data paths in different ring

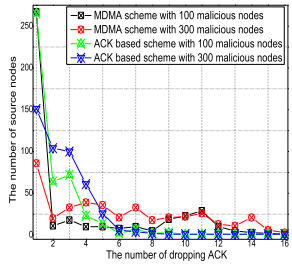


Fig. 10 The number of dropping ACK in MDMA scheme with different malicious nodes

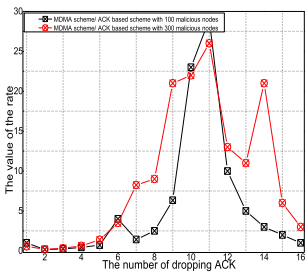


Fig. 11 The ratio of dropping data in MDMA and ACK based scheme with different malicious nodes

than in hotspots area. So choosing appropriate data paths for each node is vital for the network lifetime.

The number of the dropped ACK packet and the rate of the dropped ACK in MDMA scheme and ACK based scheme when the data packet is routed to sink successfully are respectively given in Fig. 10 and Fig. 11. It can be seen from Fig. 10 that the number of dropping ACK is less when the data is routed to sink. The more the number of the dropped ACK, the more the source node is in MDMA scheme, it can be seen from Fig. 11. Because each node in MDMA scheme produces an ACK when it receives a data packet, there are more ACK in the network, the probability which the produced ACK is dropped is bigger in MDMA scheme.

7. Conclusion

In this paper, a multi-data and multi-ACK (MDMA) scheme is proposed for detection and containing selective forwarding attacks. For detecting and resisting SFAs, it is a fundamental way to improve the effectiveness of the scheme that using multiple data routing scheme and return enough ACK to the source node. But the energy of the node in the network is limited, the previous scheme have shortcomings in the performance. MDMA scheme ingenious use the surplus energy of the non-hotspots area to build multiple data routing and ACK routing, the network lifetime can't be reduced, it has the vital significance to the development of sensor networks.

Acknowledgments

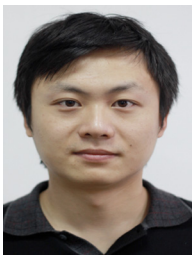
This work was supported in part by the National Natural Science Foundation of China (61379110, 61073104), The National Basic Research Program of China (973 Program) (2014CB046305).

References

[1] Y. Hu and 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.
[2] M. Dong, K. Ota, A. Liu, and M. Guo, "Joint Optimization of Lifetime and Transport Delay under Reliability Constraint Wireless Sensor Networks," IEEE Trans. Parallel Distrib. Syst., vol.27, no.1, pp.225–236, 2016.
[3] Y. Hu, M. Dong, K. Ota, et al., "Mobile Target Detection in Wireless Sensor Networks With Adjustable Sensing Frequency," IEEE System Journal, 2014.
[4] Z.K. Wazir, X. Yang, Y.A. Mohammed, and A. Quratulain, "Comprehensive study of selective forwarding attack in wireless sensor networks," International Journal of Computer Network and Information Security (IJCNIS), vol.3, no.1, p.1–10, 2011.
[5] B. Xiao, B. Yu, and C. Gao, "CHEMAS: Identify suspect nodes in selective forwarding attacks," Journal of Parallel and Distributed Computing, vol.67, no.11, pp.1218–1230, 2007.
[6] R.R. Lu, X. Lin, H. Zhu, X. Liang, and X. Shen, "BECAN: a bandwidth-efficient cooperative authentication scheme for filtering injected false data in wireless sensor networks," IEEE Trans. Parallel Distrib. Syst., vol.23, no.1, pp.32–43, 2012.
[7] X. Liu, K. Ota, A. Liu, and Z. Chen, "An incentive game based evolutionary model for crowd sensing networks," Peer-to-Peer Networking and Applications, vol.9, no.4, pp.692–711, 2016.
[8] X. Liu, M. Dong, K. Ota, P. Hung, and 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.
[9] C. Karlof and D. Wagner, "Secure routing in wireless sensor networks: attacks and countermeasures," Ad Hoc Networks, vol.1, no.2, pp.293–315, 2003.
[10] H.-M. Sun, C.-M. Chen, and Y.-C. Hsiao, "An efficient countermeasure to the selective forwarding attack in wireless sensor networks," Proc. IEEE TENCON 2007, pp.1–4, Oct. 2007.
[11] J.W. Kim, S.Y. Moon, T.H. Cho, et al., "Improved message communication scheme in selective forwarding attack detection method," 7th International Conference on Digital Content, Multimedia Technology and its Applications (IDCTA), pp.169–172, 2011.
[12] M. Dong, K. Ota, L.T. Yang, A. Liu, and M. Guo, "LSCD: A Low Storage Clone Detecting Protocol for Cyber-Physical Systems," IEEE Trans. Comput.-Aided Design Integr. Circuits Syst., vol.35, no.5, pp.712–723, 2016.
[13] Z. Zheng, A. Liu, L.X. Cai, Z. Chen, and X.S. Shen, "Energy and Memory Efficient Clone Detection in Wireless Sensor Networks," IEEE Trans. Mobile Comput., vol.15, no.5, pp.1130–1143, 2016.
[14] S. He, J. Chen, X. Li, et al., "Mobility and intruder prior information improving the barrier coverage of sparse sensor networks," IEEE Trans. Mobile Comput., vol.13, no.6, pp.1268–1282, 2014.
[15] A. Varga, The OMNET++ Discrete Event Simulation System. http://www.omnetpp.org, version 4.1.

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.

Xiao Liu received B.Sc. in 2014. Currently she is a master student with School of Information Science and Engineering of Central South University, China. Her research interest is crowd sensing networks, wireless sensor networks and wireless security.



He Li received the B.S., M.S. degrees in Computer Science and Engineering from Huazhong University of Science and Technology in 2007 and 2009, respectively. He is currently a Ph.D. candidate in Huazhong University of Science and Technology. His research interests include cloud computing and software defined networking.

Jun Long is a Professor of School of Information Science and Engineering of Central South University, China. His major research interest is wireless sensor network.