Adaptive and Channel-Aware Detection of Selective Forwarding Attacks in Wireless Sensor Networks

Ju Ren, *Student Member, IEEE*, Yaoxue Zhang, Kuan Zhang, *Student Member, IEEE*, and Xuemin Shen, *Fellow, IEEE*

***Abstract*—Wireless sensor networks (WSNs) are vulnerable to selective forwarding attacks that can maliciously drop a subset of forwarding packets to degrade network performance and jeopardize the information integrity. Meanwhile, due to the unstable wireless channel in WSNs, the packet loss rate during the communication of sensor nodes may be high and vary from time to time. It poses a great challenge to distinguish the malicious drop and normal packet loss. In this paper, we propose a channel-aware reputation system with adaptive detection threshold (CRS-A) to detect selective forwarding attacks in WSNs. The CRS-A evaluates the data forwarding behaviors of sensor nodes, according to the deviation of the monitored packet loss and the estimated normal loss. To optimize the detection accuracy of CRS-A, we theoretically derive the optimal threshold for forwarding evaluation, which is adaptive to the time-varied channel condition and the estimated attack probabilities of compromised nodes. Furthermore, an attack-tolerant data forwarding scheme is developed to collaborate with CRS-A for stimulating the forwarding cooperation of compromised nodes and improving the data delivery ratio of the network. Extensive simulation results demonstrate that CRS-A can accurately detect selective forwarding attacks and identify the compromised sensor nodes, while the attack-tolerant data forwarding scheme can significantly improve the data delivery ratio of the network.**

***Index Terms*—Wireless sensor network, selective forwarding attack, reputation system, packet dropping, channel-aware, routing.**

I. Introduction

As a promising event monitoring and data gathering technique, wireless sensor network (WSN) has been widely applied to both military and civilian applications. Many WSNs are deployed in unattended and even hostile environments to perform mission-critical tasks, such as battlefield reconnaissance and homeland security monitoring. However, due to the lack of physical protection, sensor nodes are easily compromised by adversaries, making WSN vulnerable to various security threats [1], [2]. One of the most severe threats is selective forwarding attack, where the compromised nodes can maliciously drop a subset of forwarding packets to deteriorate the data delivery ratio of the network. It also has significantly negative impacts to data integrity, especially for data-sensitive applications, e.g., health-care and industry monitoring. On the other hand, since WSNs are generally deployed in open areas (e.g., primeval forest), the unstable wireless channel and medium access collision can cause remarkable normal packet losses. The selective forwarding attacks are concealed by the normal packet losses, complicating the attack detection. Therefore, it is very challenging to detect the selective forwarding attacks and improve the network performance.

Most of related works focus on monitoring the packet losses in each transmission link and isolating the nodes with high packet loss rates from the data forwarding path [3]–[6]. These solutions can improve the data delivery ratio or network throughput but have little effect on detecting selective forwarding attacks. Since the main challenge of attack detection is to distinguish the malicious drop from normal packet loss, the normal packet loss rate of the transmission link should be considered in the forwarding evaluation. For example, a source node N_s sends 10 packets to the destination node N_d via two forwarding nodes N_a and N_b, respectively. N_a forwards 6 packets to N_d, while N_b only forwards 5 packets to N_d. Intuitively, N_a behaves better than N_b during the data forwarding. However, if the normal packet loss rates from N_s to N_a and N_b are 20% and 50%, respectively, N_a should have a higher probability to misbehave in this data forwarding. Therefore, we consider the deviation between the normal losses and actual losses as the key factor to detect selective forwarding attacks.

However, for the WSNs deployed in hostile environments where the wireless channel is unstable, normal packet loss rate highly depends on the wireless channel quality that varies spatially and temporally. If we use a measured or estimated normal packet loss rate to detect selective forwarding attacks, some innocent nodes may be falsely identified as attackers due to the time-varied channel condition. For instance, if a mobile obstacle abruptly blocks the data transmission of two sensor nodes, the unexpected packet losses may mislead the attack detection. Therefore, a flexible and fault-tolerant evaluation technique is crucial to accurately identify the attacks and compromised sensor nodes [7], [8]. Meanwhile, due to the negative impacts of selective forwarding attacks, data delivery ratio of a network becomes the primary performance metric for resisting the

Manuscript received August 20, 2015; revised January 10, 2016; accepted January 28, 2016. Date of publication February 8, 2016; date of current version May 6, 2016. This work was supported in part by the International Science and Technology Cooperation Program of China under Grant 2013DFB10070, in part by the China Hunan Provincial Science and Technology Program under Grant 2012GK4106, in part by the National Natural Science Foundation of China under Grant 61272149, and in part by the NSERC, Canada. The associate editor coordinating the review of this paper and approving it for publication was T. Taleb. *(Corresponding author: Kuan Zhang.)*

J. Ren and Y. Zhang are with the School of Information Science and Engineering, Central South University, Changsha 410083, China (e-mail: ren_ju@csu.edu.cn; zyx@csu.edu.cn).

K. Zhang and X. Shen are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: k52zhang@bbcr.uwaterloo.ca; xshen@bbcr.uwaterloo.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2016.2526601

attacks. Although compromised sensor nodes can be accurately identified, they are still available candidates to forward data for other sensor nodes before physically renewed or replaced. If a compromised node launches attack with a low probability but has good channel condition, it may forward more data packets than a normal node with poor channel condition, in spite of the malicious drops. Therefore, it is of paramount importance to design an attack-tolerant routing scheme to make full use of these nodes or stimulate their cooperation for improving the data delivery ratio.

In this paper, we propose a Channel-aware Reputation System with adaptive detection threshold (CRS-A) to detect selective forwarding attacks in WSNs. Specifically, we divide the network lifetime to a sequence of evaluation periods.[1] During each evaluation period, sensor nodes estimate the normal packet loss rates between themselves and their neighboring nodes, and adopt the estimated packet loss rates to evaluate the forwarding behaviors of its downstream neighbors along the data forwarding path. The sensor nodes misbehaving in data forwarding are punished with reduced reputation values by CRS-A. Once the reputation value of a senor node is below an alarm value, it would be identified as a compromised node by CRS-A. Compared to our previous work [10], this paper has the following enhancements and new contributions.

(i) We propose CRS-A, which evaluates the forwarding behaviors of sensor nodes by utilizing an adaptive detection threshold. By theoretically analyzing its performance, we derive an optimal detection threshold for evaluating the forwarding behaviors to optimize the detection accuracy of CRS-A. The optimal detection threshold is determined for each transmission link in a probabilistic way, and can also be adaptive to the time-varied channel condition and the attack probability of the forwarding node.

(ii) We develop a distributed and attack-tolerant data forwarding scheme to collaborate with CRS-A for stimulating the forwarding cooperation of compromised nodes and improving the data delivery ratio of the network. Rather than isolating all the compromised nodes from data forwarding, it jointly considers the time-varied channel condition and attack probabilities of neighboring nodes in choosing forwarding nodes.

(iii) Extensive simulation results demonstrate that the proposed CRS-A with attack-tolerant data forwarding scheme can achieve a high detection accuracy with both of false and missed detection probabilities close to 0, and improve more than 10% data delivery ratio for the network.

The remainder of this paper is organized as follows. Section II reviews the related works. Section III introduces the system model and design goals. The proposed CRS-A is detailed in Section IV and the adaptive detection threshold is determined in Section V. Section VI presents the attack-tolerant data forwarding and summarizes the adaptive and channel-aware forwarding evaluation scheme. Section VII validates the performance of the proposed scheme by extensive simulation results. Finally, Section VIII concludes the paper and outlines our future works.

II. Related Work

Increasing attention has been paid to developing countermeasures against selective forwarding attacks, due to their negative impacts on network performance and information integrity. The basic idea of existing works is to monitor the forwarding behaviors of sensor nodes, which can provide evidence and guidance for attack detection and defense [11]. In the following literature review, we divide the existing works into two categories: acknowledgment based and neighbor-surveillance based schemes, according to different monitoring techniques for data forwarding.

A. Acknowledgment Based Defense Techniques

This type of schemes is to use acknowledgments from different nodes in the routing path to determine the packet loss rate of each hop and detect the attackers [12], [13]. Xiao et al. [3] propose a scheme that randomly chooses a number of intermediate nodes along a forwarding path as checkpoints to return acknowledgments for each received packet. If suspicious behavior is detected, it generates an alarm packet and delivers it to the source node. Shakshuki et al. [14] design and implement an intrusion-detection system, named Enhanced Adaptive ACKonwledgment (EAACK), for mobile ad hoc networks. Due to the high load of hop-by-hop acknowledgments, EAACK combines a two-hop acknowledgment scheme and an end-to-end acknowledgment scheme to detect the malicious behaviors and reduce the network overhead. In addition, EAACK adopts a digital signature with acknowledgment to ensure authentication, integrity, and non-repudiation. As an elastic evaluation scheme, reputation system is also applied to attack detection. Zhang et al. [4] develop an audit-based misbehavior detection system to integrate reputation management, trustworthy route discovery, and identification of misbehaving nodes based on behavior audits in ad hoc networks. In [15], the correlations between link errors and malicious drops are investigated to detect selective forwarding attacks. In order to guarantee truthful calculation for the correlations, they propose a homomorphic linear authenticator (HLA) based public auditing architecture that allows the detector to verify the truthfulness of acknowledgments reported by nodes.

B. Neighbor-Surveillance Based Defense Techniques

With the Watchdog hardware [16], sensor nodes can monitor the forwarding behaviors of their neighboring nodes and record the actual packet loss accurately. Suat Ozdemir [5] investigates a functional reputation based reliable data aggregation method against selective forwarding attacks in clustered WSNs. Each node maintains a reputation table to evaluate the behaviors of its neighbor nodes, based on the forwarding monitoring of the neighboring nodes. The nodes with low reputation

[1] The network lifetime is divided into a sequence of time slots. Since we periodically evaluate the forwarding behaviors of sensor nodes in each time slot, we use "evaluation period" to replace "time slot" in this paper similar to [9].

values are isolated from the routing path. However, the reputation evaluation is only based on the monitored packet loss during the forwarding. Hao et al. [6] design a repeated game based approach to analyze the collusion on selective forwarding attacks in multi-hop wireless networks. In [17], Li et al. propose a Side Channel Monitoring (SCM) scheme to detect selective forwarding attacks in wireless ad hoc networks. SCM use the nodes adjacent to a data communication route, to constitute a side channel for monitoring the forwarding behaviors of the nodes en route. Once misbehaviors are detected, the monitoring nodes send alarm packets to the source node through both channels.

Besides these two categories of countermeasures, multi-path routing is also a widely applied technique to minimize the impact of selective forwarding attacks on data delivery rather than detect them [18]–[20]. The idea is to divide each data packet into M shares by a (T, M)-threshold secret sharing algorithm. Each packet share is assigned a TTL (time to live) field and forwarded by a randomly selected neighboring node. As the TTL decreases after each transmission, the random forwarding is repeated until TTL decreases to 0. As long as the destination receives T shares, the original message can be successfully reconstructed. In such a way, the data integrity can be guaranteed.

Most of related works discussed above can effectively mitigate the negative impacts of selective forwarding attacks on information integrity and network performance. However, they have limited capability to accurately detect the attacks and identify the compromised sensor nodes. Several recent studies consider the normal packet loss into selective forwarding attack detection for wireless mesh networks [21], [22]. However, both of the works use an estimated normal packet loss rate to evaluate the data forwarding behaviors over a long period. Such approaches are not applicable for the WSNs in unstable radio environment, where the high and time-varied packet loss may significantly reduce detection accuracy. Moreover, in their schemes, a node will be identified as an attacker once the number of lost packets during its forwarding exceeds a certain value. The one-time detection can also produce a large false detection probability for the innocent nodes [23]. In our previous work [10], a reputation system is exploited to detect selective forwarding attacks by taking the normal packet loss rate into consideration. However, it is based on a fixed evaluation threshold and simply isolates all the compromised nodes from the data forwarding paths. In this paper, we determine an adaptive threshold to evaluate the data forwarding behaviors, which can optimize the detection accuracy of the reputation system. Moreover, we develop an attack-tolerant routing scheme collaborating with the reputation system to stimulate the cooperation of compromised nodes for an improved data delivery ratio.

III. System Model and Design Goals

A. Network Model

We consider a WSN consisting of a set of randomly distributed sensor nodes, denoted by $\mathcal{N}$, and a sink node to monitor an open area. Each sensor node periodically senses the interested information from the surroundings, and transmits the sensed data to the sink via multi-hop routing among sensor nodes. Sensor nodes communicate with their neighboring nodes based on the IEEE 802.11 DCF. The monitored area has an unstable radio environment, making the packet loss rates during the communications of sensor nodes significantly increased and vary from time to time [21].

TABLE I
Frequently Used Notations

Notation	Definition
$T_t \in \mathfrak{T}$, P_M	An evaluation period, compromising probability
$S_{i,j}(t)$	The number of data packets sent from N_i to N_j during the evaluation period T_t
$m_{i,j}(t)$, $p_{i,j}(t)$	The number of data packets forwarded by N_j and the estimated normal loss rate between N_i and N_j in T_t
$r^1_{i,j}(t)$, $r^2_{i,j}(t)$	The first-hand and second-hand short-term reputation score of N_j evaluated by N_i in T_t
$R_{i,j}$, $R^I_{i,j}(t)$	The long-term reputation value of n_j in N_i's reputation table, the integrated value of $r^1_{i,j}(t)$ and $r^2_{i,j}(t)$ in T_t
$\xi_{i,j}(t)$	The misbehaving detection threshold in T_t for evaluating $m_{i,j}$ in reputation evaluation
R_s, R_m, R_a	The minimum and maximum reputation values, the alarm reputation value to identify a malicious node
$\overline{R}'_j$, NC_i	Adjusted average reputation value of N_j, the set of N_i's neighboring nodes
δ, λ, α, σ	Adjustment and punishment for reputation evaluation, penalty for calculating second-hand reputation score, weight of the first-hand reputation score for reputation integration
$\eta_{i,j}(t)$, $\mu_{i,j}(t)$	Missed and false detection probability for evaluating N_j
X, Y	Random variables denoting the number of normally lost packets from N_i to N_j, and the number of packets maliciously dropped by N_j
Y'	Random variable denoting the number of packets dropped by N_j, when N_j is malicious
p_j, $N_{i,f}(t)$	Estimated attack probability of N_j, data forwarding node of N_i in T_t

Since sensor nodes are deployed in open area and lack adequate physical protection, they may be compromised by adversaries through physical capture or software vulnerabilities to misbehave in data forwarding. We use P_M to denote the compromising probability of sensor node, which is defined as the probability that a sensor node is compromised by the adversary. Meanwhile, we assume that sensor nodes can monitor the data forwarding traffic of their neighboring nodes by neighbor monitoring with Watchdog [16] or acknowledgment-based approaches [12]. It means that a sensor node can obtain that how many data packets are forwarded by its forwarding sensor nodes. Existing works [5], [13] provide a comprehensive study on monitoring forwarding traffic of sensor nodes, which is not the focus of this paper. Since the unstable radio environment causes fluctuated packet loss rates between the neighboring nodes, it is challenging to distinguish the monitored forwarding behavior is normal or not. For easy understanding of the work, Table I summarizes the frequently used mathematical notations.

B. Attack Model

Compromised sensor nodes can launch selective forwarding attacks to degrade the performance of the network. Specifically, when a compromised sensor node receives a data packet, it maliciously drops it with a probability, referred to as attack probability. Since the adversary can control the attack probabilities of compromised nodes, it is difficult to distinguish if the packet losses are caused by fluctuated channel condition or malicious drops, especially for the nodes with low attack probabilities [24].

Furthermore, several neighboring compromised sensor nodes can collaborate with each other to launch promotion/demotion attacks to achieve benefits [25]. For example, if N_a and N_b are two neighboring compromised sensor nodes and data traffic is from N_a to N_b, N_a may provide a partial evaluation for N_b's forwarding behaviors. Besides, N_a can announce N_b as a normal node to its other neighboring nodes, in spite of N_b misbehaving in the data forwarding. However, we do not consider the special case where N_a is totally honest in data forwarding to cover for N_b's misbehaviors to achieve benefits. This case can be effectively addressed by the hop-by-hop acknowledgment or two directional neighbor monitoring techniques [4], [22].

We consider that cryptographic techniques have been utilized in the network to provide sufficient data confidentiality and authentication against the adversary, then we can focus on resisting selective forwarding attacks. In addition, we assume there are only a fraction of sensor nodes compromised by the adversary to misbehave in data forwarding, since the network would be useless if the majority of sensor nodes are manipulated by the adversary. In the following, we call the compromised sensor nodes as malicious nodes, and the other sensor nodes as normal nodes.

C. Design Goals

The objective of this paper is to detect selective forwarding attacks based on the monitored forwarding traffic information and improve the data delivery ratio for WSNs. Specifically, the proposed scheme aims to achieve the following two goals.

1) Detection accuracy: A high detection accuracy should be achieved for detecting selective forwarding attacks and identifying the malicious nodes, which can be measured by two metrics. The one is the attacks should be accurately detected once the malicious nodes misbehave in data forwarding. The other is normal nodes cannot be falsely detected as malicious nodes due to the fluctuated normal packet losses.

2) Data delivery ratio improvement: Besides the detection of selective forwarding attacks, the data delivery ratio of the network should be improved by the proposed scheme to mitigate the negative impacts caused by the attacks. Meanwhile, the proposed scheme should be able to partly stimulate the cooperation of malicious nodes in data forwarding.

IV. CRS-A: The Channel-Aware Reputation System With Adaptive Detection Threshold

In this section, we propose CRS-A to detect selective forwarding attacks and identify malicious nodes. In CRS-A, each sensor node maintains a reputation table to evaluate the long-term forwarding behaviors of its neighboring nodes. The essence of CRS-A is to dynamically update the reputation table based on the forwarding behavior evaluation for the neighboring nodes, by taking the normal packet loss rate into consideration. However, as the unstable radio environment make the quality of wireless channel vary with time, normal packet loss may be different over a long time period. Therefore, we divide the whole network lifetime into a sequence of evaluation periods $\mathcal{T} = \{T_1, \ldots, T_t, \ldots\}$. In each evaluation period T_t, the channel condition of each data transmission link is assumed to be stable. Meanwhile, for each T_t, we introduce a channel estimation stage at the beginning of T_t, and a reputation update stage at the end of T_t. During the channel estimation stage, sensor nodes estimate the normal packet loss rates of the communication links with their neighboring nodes, and use them to evaluate the forwarding behaviors of neighboring nodes. Fig. 1 shows the overview of evaluation periods over the network lifetime.

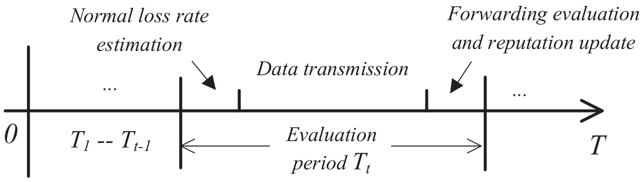
Normal loss rate estimation
Forwarding evaluation and reputation update
Data transmission
...
...
0
T_1 -- T_{t-1}
Evaluation period T_t
T

Fig. 1. The overview of evaluation periods.

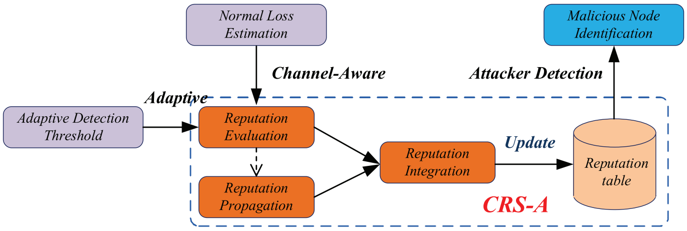

Fig. 2. The architecture of CRS-A.

The reputation update in CRS-A consists of three procedures: reputation evaluation, propagation and integration. *Reputation Evaluation* is to evaluate short-term reputation scores for the forwarding behaviors of sensor nodes, based on the deviation of estimated normal packet loss rate and monitored actual packet loss rate. With *Reputation Propagation*, the evaluated short-term reputation scores can be propagated within the neighboring nodes to achieve a more comprehensive evaluation. Finally, by *Reputation Integration*, sensor nodes integrate the reputation scores evaluated by themselves and the propagated reputation scores from their neighboring nodes to update the reputation table. Fig. 2 shows the architecture of CRS-A. In the following, we describe each procedure of CRS-A in detail.

A. Normal Packet Loss Estimation

Since the wireless channel of the WSN is easily impacted by unstable radio environment to cause noticeable packet losses during wireless transmission, the normal packet loss should be considered into the forwarding behavior evaluation for sensor nodes. According to the network model, normal packet loss is mainly caused by the poor and unstable wireless channel and MAC layer collisions. We discuss the normal packet loss estimation from the two aspects as follows.

1) Packet Loss Caused by Radio Link Quality: The poor and unstable radio link quality is the primary reason for the time-varied packet losses. In [21], [22], the link condition is formulated as a two-state Markov model, and the packet loss rate is determined as an average value over a long-term period. However, adopting an average value to represent a time-varied value may mislead the evaluation for forwarding behaviors [26], [27]. Furthermore, dynamic environments make the link quality varied in different locations. Therefore, the packet loss estimation should be performed in each evaluation period by each sensor node. In CRS-A, the link quality estimation for each pair of neighboring nodes is based on the Received Signal Strength Indicator (RSSI) and Signal-to-Noise Ratio (SNR), under the symmetric channel assumption [27], [28]. For each T_t, the packet loss rate caused by poor link quality, denoted by $p^1_{i,j}(t)$, can be estimated by RSSI and SNR for the transmission link from N_i to N_j.

2) Packet Loss Caused by MAC Layer Collisions: As data transmission between two neighboring nodes is based on the IEEE 802.11 DCF, MAC layer collisions may increase the normal packet loss rate. Since sensor nodes are static in our network, it means each sensor node has a fixed number of neighboring nodes. Then, we can use the analytical results in [21], [29] to estimate the packet loss caused by medium access collisions without the impact of hidden terminals [26], [27]. Let n be the number of nodes contending for channel access at N_j and p_t as the probability that a node transmits data in time slot. When MAC channel is at steady state, the probabilities for observing an idle, successful, and colliding slot, denoted as p_i, p_s, and p_c, respectively, are

$$\begin{cases} p_i = (1-p_t)^n \\ p_s = n \cdot p_t \cdot (1-p_t)^{n-1} \\ p_c = 1 - p_i - p_s. \end{cases} \tag{1}$$

And the channel busy ratio R_b can be calculated as

$$C_b = 1 - (p_i \cdot t_d)/(p_i \cdot \sigma + p_s \cdot t_s + p_c \cdot t_c), \tag{2}$$

where t_d, t_s and t_c denote the idle slot length, the duration of a successful transmission, and the duration of a collision, respectively, which can be determined by [30].

Therefore, the packet loss rate caused by MAC layer collisions $p^2_{i,j}$ is the probability that a node encounters collisions when it transmits, i.e.,

$$p^2_{i,j} = 1 - (1-p_t)^{n-1}. \tag{3}$$

Combining Eq. (1) and (2), C_b can be expressed as a function of p_t. Since C_b can be obtained by channel monitoring, p_t can be calculated to determine $p^2_{i,j}$ based on Eq. (3).

In summary, the estimated normal packet loss rate between N_i and N_j in T_t is $p_{i,j}(t) = p^1_{i,j}(t) + p^2_{i,j} - p^1_{i,j}(t)p^2_{i,j} \approx p^1_{i,j}(t) + p^2_{i,j}$.

B. Reputation Evaluation

In CRS-A, sensor nodes monitor their neighbors to evaluate reputation scores for their forwarding behaviors during each evaluation period. The evaluated reputation scores are named as first-hand reputation scores. Specifically, in the data transmission stage of T_t, node N_i ($N_i \in \mathcal{N}$) records the number of data packets sent to its next hop node N_j as $S_{i,j}(t)$, and the number of data packets forwarded by N_j as $f_{i,j}(t)$. Thus, the number of data packets lost in the transmission from N_i to N_j is $m_{i,j}(t) = S_{i,j}(t) - f_{i,j}(t)$. Based on the discussion of the previous subsection, we can estimate the normal packet loss rate between N_i and N_j as $p_{i,j}(t)$. Since each data packet is transmitted to N_j independently, the data transmission from N_i to N_j can be regarded as a sequence of independent repeated trials. It means, if N_i sends l data packets to N_j, the probability of k ($0 \le k \le l$) out of l packets lost during the transmission, denoted by $P_{i,j}(X = k)$, follows a binomial distribution, i.e.,

$$P_{i,j}(X = k) = \binom{l}{k} (p_{i,j}(t))^k (1 - p_{i,j}(t))^{l-k}. \tag{4}$$

We consider the forwarding behavior evaluation for N_j during an evaluation period T_t as a sampling test. If N_j behaves normally during data forwarding, $m_{i,j}(t)$ should slightly fluctuate around the estimated number of normal lost data packets $p_{i,j}(t) \cdot S_{i,j}(t)$. However, when $m_{i,j}(t) > p_{i,j}(t) \cdot S_{i,j}(t)$, with the increase of $m_{i,j}(t)$, the probability of N_j misbehaving in data forwarding increases. In order to evaluate $m_{i,j}(t)$, we introduce a detection threshold $\xi_{i,j}(t)$ ($S_{i,j}(t) \cdot p_{i,j}(t) < \xi_{i,j}(t) < S_{i,j}(t)$, $\xi_{i,j}(t) \in \mathbb{N}^+$) and define the reputation evaluation function of N_i to N_j as follows.

$$r^1_{i,j}(t) = \begin{cases} +\delta, & \text{if } m_{i,j}(t) \le p_{i,j}(t) \cdot S_{i,j}(t) \\ -\delta, & \text{if } p_{i,j}(t) \cdot S_{i,j}(t) < m_{i,j}(t) \le \xi_{i,j}(t) \\ -\lambda, & \text{if } m_{i,j}(t) > \xi_{i,j}(t). \end{cases} \tag{5}$$

where λ is a punishment factor and δ is a adjustment factor. We set $\lambda \gg \delta$ and explain the function as follows.

- If $m_{i,j}(t) \le p_{i,j}(t) \cdot S_{i,j}(t)$, the sampling test is acceptable, which means the transmission between N_i and N_j is successful. Thus, N_i rewards a positive δ to N_j.
- If $p_{i,j}(t) \cdot S_{i,j}(t) < m_{i,j}(t) \le \xi_{i,j}(t)$, we consider it is a normal fluctuation of $p^m_{i,j}$ around $p_{i,j}$, and rate $-\delta$ to N_j to neutralize the reputation evaluation.
- When $m_{i,j}(t) > \xi_{i,j}(t)$, we consider there is a high probability for N_j to misbehave in the data forwarding. If it happens, N_i rates a punishment $-\lambda$ to N_j.

As we discussed above, if N_j is a normal node, $m_{i,j}(t)$ will slightly fluctuate around $p_{i,j}(t) \cdot S_{i,j}(t)$. The proposed reputation evaluation function should make the reputation value of N_j stable or increased after a number of evaluation periods. On the other hand, if N_j misbehaves in data forwarding, $m_{i,j}(t)$ may be larger than $p_{i,j}(t) \cdot S_{i,j}(t)$ with a high probability. The proposed function should decrease the reputation value of N_j sharply after a number of evaluation periods. According to Eq. (5), it can be found that both of the expected two characteristics are highly impacted by the value of $\xi_{i,j}(t)$. Therefore, how to determine the optimal $\xi_{i,j}(t)$ is of significant importance for improving the performance of the reputation evaluation, which will be discussed in the next section.

C. Reputation Propagation

In order to share the monitored forwarding behavior information and hence to improve the attack detection accuracy, N_i propagates the first-hand reputation scores, such as $r^1_{i,j}(t)$, to their neighbors during each T_t. The received reputation scores from the neighboring nodes are called as second-hand reputation scores, which reflect the evaluation of the neighboring nodes on their next hop nodes. However, the reputation propagation causes CRS-A vulnerable to collaborative promotion/demotion attacks, which means neighboring malicious nodes can collaborate with each other to mutually promote their reputation scores [25]. To mitigate the impact of the potentially partial reputation scores, we determine the second-hand reputation scores as follows.

Denote the set of N_i's neighboring sensor nodes as NC_i, and the number of nodes in NC_i as $|NC_i|$. We further divide the nodes of NC_i into two subsets, $NC_{i,g}$ and $NC_{i,b}$, based on their long-term reputation values in N_i. Let N_s be a node of NC_i. We put N_s into the honest neighbor set $NC_{i,g}$, if $R_{i,s} > \frac{\sum_{x \in NC_i} R_{i,x}}{|NC_i|}$. Otherwise, N_s is allocated to the dishonest neighbor set $NC_{i,b}$. Therefore, we calculate the second-hand reputation score of N_i to its neighboring node N_j as

$$r^2_{i,j}(t) = \sum_{x \in NC_{i,g}} \frac{R_{i,x}}{\sum_{s \in NC_i} R_{i,s}} \cdot r^1_{x,j}(t) + \sum_{x \in NC_{i,b}} \frac{R_{i,x}}{\sum_{s \in NC_i} R_{i,s}} \cdot \alpha r^1_{x,j}(t) \qquad (6)$$

where α is a penalty factor to reduce the weight of the information propagated by the potentially dishonest neighbors and $\alpha < 1$.

Since the long-term reputation values of malicious nodes may decrease after misbehaving in a number of evaluation periods, these nodes are classified into the dishonest neighbor set and the weights of their propagating information are reduced by the penalty factor α. As a result, the negative impacts of mutual reputation promotions among neighboring malicious nodes can be significantly mitigated by Eq. (6). To reduce the communication overhead of reputation propagation, the propagated reputation scores can be piggybacked to other data packets, such as the periodically exchanged neighbor information.

D. Reputation Integration

After reputation propagation, the first-hand and second-hand short-term reputation scores should be integrated to update the reputation table. Denote $R_{i,j}$ as the long-term reputation value of N_j in N_i's reputation table, and R_m and R_s as the upper bound and lower bound of reputation value. We calculate the integrated reputation score as $R^I_{i,j}(t) = \sigma r^1_{i,j}(t) + (1 - \sigma) r^2_{i,j}(t)$, and update $R_{i,j}$ as the following equation.

$$R_{i,j} = \begin{cases} R_s, & if\, R_{i,j} + R^I_{i,j} \le R_s \\ R_{i,j} + R^I_{i,j}, & if\, R_s < R_{i,j} + R^I_{i,j} < R_m \\ R_m, & otherwise. \end{cases} \qquad (7)$$

Here, σ is the weight factor of the first-hand information and $\sigma > 0.5$. R_m and R_s are system parameters that can be chosen based on the system requirements. For instance, we can set $R_s = 0$ and $R_m = 255$ to keep each reputation value only take 1 byte. Such that, the storage requirement of the sensor nodes and the communication overhead can be reduced.

E. Malicious Nodes Identification

In each T_t, sensor nodes can evaluate the forwarding behaviors of their next hop sensor nodes and update their reputation table with the above three procedures. After a number of evaluation periods, the reputation values of malicious nodes are significantly reduced in the reputation tables of their neighboring nodes. To identify the malicious nodes, sensor nodes send their reputation tables to the sink for identification after a fixed time. When the average reputation value in N_j's neighbors is below R_a, i.e., $\frac{\sum_{N_i \in NC_j} R_{i,j}}{|NC_j|} < R_a$, N_j is identified as a malicious node. Here, R_a is an alarm reputation value that can be predefined according to system requirements. If N_j is identified as a malicious node, the network operator can perform a security check or software reset for these nodes. However, since malicious nodes can mutually promote their reputation values or collaboratively degrade the reputation values of normal nodes, the average reputation value should be adjusted against the promotion and demotion attacks [31].

We denote the original average reputation value of N_j as $\overline{R}_j = \frac{\sum_{N_i \in NC_j} R_{i,j}}{|NC_j|}$, and the adjusted average reputation value as $\overline{R}'_j$. Then, the standard deviation $D_s(N_j)$ of all the reputation values $R_{i,j}$ ($N_i \in |NC_j|$) is

$$D_s(N_j) = \sqrt{\frac{1}{|NC_j|} \sum_{N_i \in \{NC_j\}} (R_{i,j} - \overline{R}_j)^2}. \qquad (8)$$

We define a standard deviation threshold as d_h [31], [32]. If $D_s(N^i_k) \le d_h$, we have $\overline{R}^d_k = \overline{R}_k$. Otherwise, it is suspicious that there is a promotion or demotion attack. We remove the $R_{i,j}$ with largest deviation to $\overline{R}_j$ and recalculate the $D_s(N_j)$ until we have $D_s(N_j) \le d_h$. Then, we use the average value of the rest $R_{i,j}$ as $\overline{R}'_j$. If $\overline{R}'_j < R_a$, N_j will be identified as a malicious node.

V. Adaptive Detection Threshold for CRS-A

As we discussed in Sec. IV-B, the detection accuracy of CRS-A is significantly impacted by the misbehaving detection threshold for reputation evaluation. In this section, we aim to determine the optimal evaluation threshold for each pair of neighboring nodes along the data forwarding path to optimize the detection accuracy of CRS-A. According to the attack model, malicious nodes can launch attacks with different probabilities, which indicates the detection threshold should be different for each communication link. Meanwhile, due to the nature of dynamic routing and time-varied channel condition in WSNs, the detection threshold should be adaptive to the time-varied data traffic and normal packet loss rate of the link. Without loss of generality, we focus on determining the optimal threshold for the transmission from N_i to N_j during the period T_t, in the following analysis.

A. Metrics of Detection Accuracy

Since CRS-A is proposed to detect selective forwarding attacks and identify malicious nodes, we first identify some performance metrics to evaluate CRS-A before optimizing them. According to Eq. (5), if $\xi_{i,j}(t)$ is set as a large value, the forwarding misbehavior of N_j will be regarded as a normal fluctuation, without being punished with $-\lambda$. It means the attacks launched by N_j are not detected by the detection of CRS-A. On the other hand, if $\xi_{i,j}(t)$ is set as a small value close to $S_{i,j}(t) \cdot p_{i,j}(t)$, the normal fluctuation of $m_{i,j}(t)$ will be detected as a misbehavior, when N_j acts normally in data forwarding. It leads to that a normal sensor node has a large probability to be falsely identified as a compromised node by the detection of CRS-A. Therefore, there exists a trade-off in determining the value of $\xi_{i,j}(t)$ to optimize the detection accuracy for selective forwarding attacks.

To this end, we introduce two metrics, missed detection probability and false detection probability. The *Missed Detection Probability* is the probability that a malicious forwarding behavior is detected as a normal behavior, while the *False Detection Probability* refers to the probability that a normal forwarding behavior is detected as a malicious behavior. If we use X to denote the data packets lost in the transmission from N_i to N_j, and Y to denote the data packets maliciously dropped by N_j, the missed detection probability $\eta_{i,j}(t)$ is

$$\begin{aligned}\eta_{i,j}(t) &= P\{X+Y \leq \xi_{i,j}(t) | j \text{ misbehaved in } T_t\}\\ &= P\{X+Y \leq \xi_{i,j}(t) | Y > 0\}\\ &= \frac{P\{\{X+Y \leq \xi_{i,j}(t)\} \cap \{Y > 0\}\}}{P\{Y>0\}},\end{aligned} \tag{9}$$

and the false detection probability $\mu_{i,j}(t)$ is

$$\begin{aligned}\mu_{i,j}(t) &= P\{X+Y > \xi_{i,j}(t) | j \text{ behaved well in } T_t\}\\ &= P\{X+Y > \xi_{i,j}(t) | Y = 0\}\\ &= P\{X > \xi_{i,j}(t)\}.\end{aligned} \tag{10}$$

Since both X and Y are discrete random variables, the probability mass function (PMF) of X and Y should be determined for calculating $\eta_{i,j}(t)$ and $\mu_{i,j}(t)$. As X is defined as the number of normally lost data packets during the transmission, the PMF of X should be Eq. (4). If the number of data packets sent by N_i during T_t is $S_{i,j}(t)$, the false detection probability $\mu_{i,j}(t)$ is the CDF of X, i.e.,

$$\mu_{i,j}(t) = 1 - \sum_{k=0}^{\xi_{i,j}(t)} \left[\binom{S_{i,j}(t)}{k} (p_{i,j}(t))^k (1-p_{i,j}(t))^{S_{i,j}(t)-k} \right]. \tag{11}$$

However, due to $\eta_{i,j}(t)$ depending on the variable Y, we should determine the PMF of Y and $X+Y$. According to the attack model, each sensor nodes has a probability P_M to be compromised by the adversary. It means $P\{Y=0\} = 1 - P_M$ and $P\{Y = Y'\} = P_M$, where Y' is a discrete random variable denoting the number of maliciously dropped packets by N_j when N_j is a malicious node.

According to the attack model, when a malicious node successfully receives a data packet, it decides to maliciously drop the packet with a probability, which is called attack probability. We denote the attack probability of N_j as p_j. Since the number of data packets sent by N_i during the evaluation t are $S_{i,j}(t)$, the PMF of Y' should be a binomial function with the number of experiments as $A_i(t) = S_{i,j}(t) - X$. Obviously, $A_i(t)$ is a random variable depending on X, so we first calculate the conditional probability when $A_i(t)$ is fixed as a, $(0 \leq a \leq S_{i,j}(t)$, $0 \leq k \leq a)$ as

$$P\{Y' = k | A_i(t) = a\} = \binom{a}{k} p_j^k (1-p_j)^{a-k}. \tag{12}$$

And the PMF of Y' is

$$\begin{aligned}P\{Y'=k\} &= \sum_{a=0}^{S_{i,j}(t)} \left[P\{Y'=k|A_i(t)=a\} P\{A_i(t)=a\} \right]\\ &= \sum_{a=0}^{S_{i,j}(t)} \left[P\{Y'=k|X=S_{i,j}(t)-a\} P\{X=S_{i,j}(t)-a\} \right]\\ &= \sum_{x=0}^{S_{i,j}(t)} \left[P\{Y'=k|X=x\} P\{X=x\} \right]\\ &= \sum_{x=0}^{S_{i,j}(t)} \left[\binom{S_{i,j}(t)-x}{k} p_j^k (1-p_j)^{S_{i,j}(t)-x-k} P\{X=x\} \right]\end{aligned} \tag{13}$$

where the third step is based on substituting $S_{i,j}(t) - a$ by x. And we have $0 \leq x \leq S_{i,j}(t)$ due to $0 \leq a \leq S_{i,j}(t)$.

Therefore, we can use the PMF of Y' to determine the PMF of Y as

$$P\{Y=k\} = \begin{cases} (1-P_M) + P_M \cdot P\{Y'=0\}, & \text{if } k=0 \\ P_M \cdot P\{Y'=k\}, & \text{if } 1 \leq k \leq S_{i,j}(t) \end{cases} \tag{14}$$

According to Eq. (14), we calculate the missed detection probability $\eta_{i,j}(t)$ in Theorem 1.

Theorem 1: If N_i sends $S_{i,j}(t)$ data packets to N_j during the evaluation period T_t and the detection threshold is $\xi_{i,j}(t)$ $\left(S_{i,j}(t) \cdot p_{i,j}(t) < \xi_{i,j}(t) < S_{i,j}(t)\right)$, the missed detection probability for evaluating N_j is

$$\eta_{i,j}(t) = \frac{\sum_{k=1}^{\xi_{i,j}(t)} \left[P\{X \leq \xi_{i,j}(t) - k\} \cdot P\{Y=k\} \right]}{P_M - P_M \cdot (1-p_j)^{S_{i,j}(t)}}, \tag{15}$$

where $P\{X \leq k\}$ is the CDF of X, which can be calculated by Eq. (4), and $P\{Y=k\}$ can be calculated by combining Eq. (13) and (14).

Proof: Based on the PMF of X and Y, we further expand Eq. (9) to calculate η_j as follows.

$$\begin{aligned}\eta_j &= \frac{P\{\{X+Y \le \xi_{i,j}\} \cap \{Y > 0\}\}}{P\{Y > 0\}}\\ &= \frac{P\{\{X+Y \le \xi_{i,j}\} \cap \{\sum_{k=1}^{S_{i,j}(t)}\{Y = k\}\}\}}{P\{\sum_{k=1}^{S_{i,j}(t)} Y = k\}}\\ &= \frac{\sum_{k=1}^{S_{i,j}(t)} P\{\{X+Y \le \xi_{i,j}\} \cap \{Y = k\}\}}{\sum_{k=1}^{S_{i,j}(t)} P\{Y = k\}}\\ &= \frac{\sum_{k=1}^{S_{i,j}(t)} \left[P\{X+Y \le \xi_{i,j} | Y = k\} \cdot P\{Y = k\}\right]}{\sum_{k=1}^{S_{i,j}(t)} P\{Y = k\}}\\ &= \frac{\sum_{k=1}^{\xi_{i,j}} \left[P\{X \le \xi_{i,j} - k\} \cdot P\{Y = k\}\right]}{1 - P\{Y = 0\}}\\ &= \frac{\sum_{k=1}^{\xi_{i,j}} \left[P\{X \le \xi_{i,j} - k\} \cdot P\{Y = k\}\right]}{P_M - P_M \cdot (1 - p_j)^{S_{i,j}(t)}},\end{aligned}$$

where $P\{X \le k\}$ is the CDF of X, which can be calculated by Eq. (4), and $P\{Y = k\}$ can be calculated by combining Eq. (13) and (14). ■

B. Determining the Optimal Threshold

In this section, we determine the optimal threshold $\xi^*_{i,j}(t)$ for reputation evaluation in CRS-A. According to Eq. (15), the missed detection probability η_j depends on the attack probability of N_j (i.e., p_j). Generally, the attack probabilities of malicious nodes are various and not known by the system in advance. However, we can use the historical data to estimate p_j for each malicious node N_j. Specifically, in each T_t, N_i can estimate p_j with the following Eq. (16), i.e.,

$$p_j = \left[\frac{\sum_{w=0}^{t}[m_{i,j}(w) - S_{i,j}(w) \cdot \big(1 - p_{i,j}(w)\big)]}{\sum_{w=0}^{t}[S_{i,j}(w) \cdot \big(1 - p_{i,j}(w)\big)]}\right]^{+}, \quad (16)$$

where $[a]^+ = a$, if $a \ge 0$; otherwise, $[a]^+ = 0$.

In Eq. (16), $S_{i,j}(w) \cdot \big(1 - p_{i,j}(w)\big)$ is the expected number of forwarded data packets at time period w, while $m_{i,j}(w) - S_{i,j}(w) \cdot \big(1 - p_{i,j}(w)\big)$ is deviation between the actual number of forwarded data packets and the expected number of forwarded data packets at time period w. Thus, Eq. (16) can partially show the probability that node j attacks (or maliciously drops) in data forwarding. When p_j is small or equal to 0, we consider N_j behaves well during the past data forwarding. The false detection probability μ_j should be minimized for CRS-A. As p_j keeps increasing, N_j has an increasing probability to be an attack. It indicates that the missed detection probability η_j should be emphasized to optimize the performance of CRS-A. Meanwhile, both of the missed detection probability η_j and false detection probability μ_j depend on $\xi_{i,j}(t)$. When $\xi_{i,j}(t)$ increases, η_j increases and μ_j decreases. And if $\xi_{i,j}(t)$ decreases, the situation reverses. It means η_j and μ_j are two contradictory optimization objectives. In order to find a trade-off between them, we can integrate η_j and μ_j as a single objective function ν_j by weighting them with p_j and $1 - p_j$, respectively. The objective function is defined as $\nu_j = p_j \cdot \eta_j + (1 - p_j) \cdot \mu_j$. Therefore, for each transmission from N_i to N_j in T_t, the optimal threshold determination problem can be formulated as calculating $\xi^*_{i,j}(t)$ to

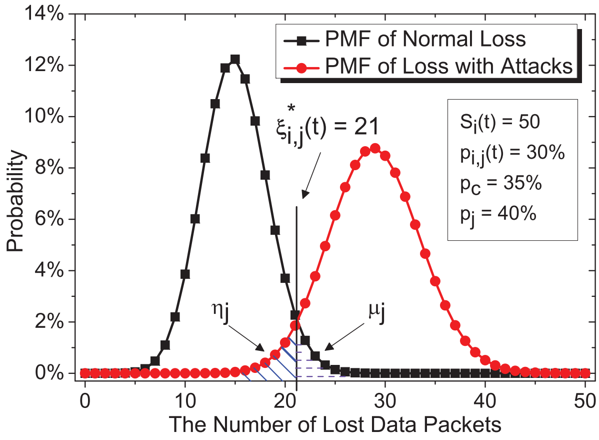

Fig. 3. An example of the optimal threshold.

$$\textbf{(PP)}\ \text{minimize}\ \nu_j = p_j \cdot \eta_{i,j}(t) + (1 - p_j) \cdot \mu_{i,j}(t)$$
$$\text{s.t.} \begin{cases} p_{i,j}(t) \cdot S_i(t) < \xi_{i,j}(t) < S_i(t) \\ \xi_{i,j}(t) \in \mathbb{N}^+ \end{cases}.$$

It is obvious that **(PP)** has only one optimization variable and a closed-form objective function. As $\xi_{i,j}(t)$ is discrete, the objective function is non-differentiable with respect to $\xi_{i,j}(t)$, which indicates the hardness of deriving a closed-form optimal solution for **(PP)**. However, due to the constraint that $\xi_{i,j}(t)$ should be an integer between $p_{i,j}(t) \cdot S_i(t)$ and $S_i(t)$, we can adopt a brute-force algorithm to calculate all the possible values for determining the optimal one. Since $S_i(t)$ is the only input variable of **(PP)** which impacts the time complexity of finding a solution, the brute-force algorithm can guarantee the time complexity is $O(S_i(t))$, i.e., $O(n)$.

Fig. 3 shows an example of the optimal threshold for N_i to evaluate N_j's forwarding behavior, where $S_i(t) = 50$, $p_{i,j}(t) = 30\%$, $P_M = 35\%$ and $p_j = 40\%$. The black line shows the PMF of the number of lost data packets, if N_j behaves normally during the forwarding. And the red line is PMF of the number of lost data packets when N_j misbehaves in forwarding and $p_j = 40\%$. Actually, the PMF of normal packet loss is $P\{X = k\}$ $(0 \le k \le S_i(t))$, while the PMF of packet loss with attacks is $P\{X + Y = k | Y > 0\}$ $(0 \le k \le S_i(t))$ that can be calculated according to Eq. (4) and (14). As shown in the figure, the optimal evaluation threshold is calculated as $\xi^*_{i,j}(t) = 21$ by solving **(PP)**. And the area of η_j and μ_j are the missed detection probability and false detection probability of the evaluation, respectively.

VI. CRS-A With Attack-tolerant Data Forwarding

As a trust evaluation technique independent of route decision, CRS-A can be applied with any data forwarding protocol for WSNs. However, due to the negative impacts of selective

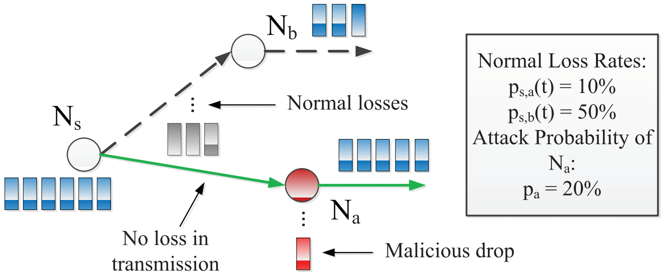

Fig. 4. An example of dynamic routing.

forwarding attacks on data forwarding, data delivery ratio is a key performance metric for evaluating a defense technique, besides the detection accuracy for attacks and malicious nodes. In this section, we first develop a distributed and attack-tolerant data forwarding scheme to collaborate with CRS-A to improve the data delivery ratio of the network. Then, we summarize the main idea and procedures of CRS-A with attack-tolerant data forwarding into an algorithm, and have a discussion on its cooperation stimulation and overhead.

A. Attack-Tolerant Data Forwarding

For a distributed data forwarding scheme, the key challenge is to decide which sensor node should be chosen in the forwarding path to optimize the network performance, based on the local knowledge [33], [34]. In this paper, we consider data delivery ratio as the primary metric of network performance. Although we can detect the malicious nodes by CRS-A, it is unreasonable to isolate all the malicious nodes from the data forwarding path. We can illustrate it with the following Fig. 4. N_a and N_b are two routing candidates of N_s, and N_a is identified as a malicious node by N_s. During T_t, N_s estimates the normal loss rate of each link as $p_{s,a}(t) = 10\%$ and $p_{s,b}(t) = 50\%$. The attack probabilities of N_a and N_b are $p_a = 20\%$ and $p_b = 0$, respectively. In this case, N_s has 6 data packets to forward. If N_s chooses N_b as the next hop, the expected number of data packets that are successfully forwarded by N_b is 3. Contrastively, the expected number of data packets forwarded by N_a should be 5, even if its reputation in N_s is low and it has an attack probability 20% according to the historical records.

To select a better forwarding node to improve the data delivery ratio, we introduce the expected data forwarding ratio (DFR), which is defined as the ratio between the expected number of forwarded data packets and the total number of sent data packets. In each evaluation period T_t, N_i chooses the node with the highest DFR from its forwarding candidate set as the next hop. The forwarding candidate set of N_i is the set of its neighboring nodes that are geographically closer to the sink than N_i. Specifically, the forwarding decision can be formulated as follows. For each N_i, given the number of data packets that N_i transmits in T_t as $S_i(t)$, if choosing N_j as the data forwarding node, the expected number of lost data packets should be $L_j(t) = S_i(t) \cdot p_{i,j}(t) + [S_i(t) - S_i(t) \cdot p_{i,j}(t)] \cdot p_j(t)$. And, the DFR of N_j is

$$\begin{aligned} DFR_j(t) &= (S_i(t) - L_j(t))/S_i(t) \\ &= 1 - p_{i,j}(t) - p_j(t) + p_{i,j}(t) \cdot p_j(t). \end{aligned} \tag{17}$$

Algorithm 1. Adaptive and Channel-aware Forwarding Evaluation during Each Evaluation Period

```
Description: Updating the reputation of sensor nodes and data
             forwarding during $T_t$ ($T_t \in \mathcal{T}$).
Phase I  Normal Loss Estimation;
for each $N_i \in \mathcal{N}$ do
    Estimate the normal packet loss rate $p_{i,j}(t)$ between $N_i$
    and each $N_j$ in $N_i$'s neighbor set;
end
Phase II  Data Transmission and Monitoring;
for each $N_i \in \mathcal{N}$ do
    Choosing $N_j$ from $RC_i$ as the next hop according to
    Eq. (17) and (18), and use $N_j$ to forward its data;
    Record the number of sent data packets $S_{i,j}(t)$ and the
    number of data packets $m_{i,j}(t)$ forwarded by $N_j$
end
Phase III  Reputation Evaluation and Updating;
for each $N_i \in \mathcal{N}$ do
    Calculate the attack probability $p_j$ of $N_j$ according to
    Eq. (16);
    Determine the optimal detection threshold $\xi^*_{i,j}(t)$ by
    solving the problem (PP);
    Evaluate the first-hand reputation score $r^1_{i,j}(t)$ according
    to Eq. (5);
    Propagate $r^1_{i,j}(t)$ to its neighboring nodes;
    if receive propagated reputation scores then
        Calculate the second-hand reputation score $r^2_{i,j}(t)$
        based on Eq. (6);
    end
    Calculate the integrated reputation score $R^I_{i,j}(t)$ with
    $r^1_{i,j}(t)$ and $r^2_{i,j}(t)$ and use it to update $R_{i,j}$ according to
    Eq. (7);
end
```

Let RC_i denote the forwarding candidate set of N_i. If N_i can directly communicate with the sink, the next hop is the sink; otherwise, N_i's next hop node $N_{i,f}(t)$ in T_t is

$$N_{i,f}(t) = \operatorname*{argmax}_{N_j \in RC_i} DFR_j(t). \tag{18}$$

B. CRS-A With Attack-tolerant Data Forwarding

Based on the preceding description on CRS-A and the attack-tolerant data forwarding, we summarize the procedures of reputation updating in CRS-A with attack-tolerant data forwarding in Algorithm 1. In the following, we discuss how the CRS-A with attack-tolerant data forwarding stimulates malicious nodes to cooperate during data forwarding, and the overhead of maintaining CRS-A in the network.

1) Cooperation Stimulation: According to Algorithm 1, when a malicious node N_j is selected into the routing path by N_i, the evaluation threshold is determined by $p_{i,j}$ and p_j to evaluate its forwarding behavior in the current evaluation period. If N_j misbehaves in this period with a probability p'_j that is higher than p_j, i.e., $p'_j > p_j$, the number of lost data packets will be larger than the evaluation threshold and it will be punished with a negative reputation score. Only if N_j adopts a lower attack probability, it could avoid a reputation punishment. For the irrational malicious nodes increasing the attack probability without considering the punishment, they are removed by the security check soon. Meanwhile, rational malicious nodes can be stimulated to behave better to achieve an improved data delivery ratio.

2) Overhead: We consider the overhead of maintaining CRS-A, in terms of its storage overhead and communication overhead. In CRS-A, each node maintains a reputation table to record the reputation values of its neighboring nodes, which produces the storage overhead for sensor nodes. If the range of reputation value is set as $[0, 255]$, each reputation value only take 8 bits and the total storage overhead of N_i for maintaining the CRS-A is $8 \cdot |NC_i|$ bits, where NC_i is the neighbor set of N_i. The communication overhead of CRS-A is mainly produced by channel estimation and reputation propagation [2]. Let B be the number of bits in a PROBE packet that sensor nodes broadcast to their neighboring nodes for channel estimation [28]. The overhead for channel estimation is B bits data broadcasting and $B \cdot |NC_i|$ bits data receiving for each node in an evaluation period. Similarly, each sensor node evaluates a reputation score for its data forwarding node, and propagates the score to its neighboring nodes in each evaluation period. Thus, the communication overhead of reputation propagation includes 8 bits data broadcasting and $8 \cdot |NC_i|$ bits data receiving. Since the PROBE packet and reputation score information are much smaller than the transmitted data packets of sensor nodes, it means CRS-A has a small communication overhead to be employed into WSNs.

VII. Simulation Results

In this section, we evaluate the performance of CRS-A and the attack-tolerant routing scheme by the simulations on OMNET++ [33], [35]. The simulation scenario consists of 100 stationary sensor nodes uniformly distributed in a $500m \times 500m$ area. The sink node is located at the center of the area. Each sensor node has a probability P_M to be compromised as a malicious node, the value of which is identified in different simulations. The attack probability of each malicious nodes N_j is randomly initialized as a value $p_j \in [0.1, 0.6]$. Each sensor node generates 10 data packets to transmit to the sink via multi-hop routing in each evaluation period, and the transmission range of a sensor node is 85m. The communication between two neighboring node is based on the IEEE 802.11 DCF, while a finite state Markov model is adopted to model the unstable wireless channel [21], [22]. Although data delivery ratio can be improved by data retransmission, we assume no retransmission technique is applied in the simulation, where we can only focus on the impacts of selective forwarding attacks on data integrity.

We setup the parameters of CRS-A as follows. The range of the reputation value of a sensor nodes is $[0, 200]$, i.e., $R_s = 0$ and $R_m = 200$. The initial reputation is 100 for all the sensor nodes. The value of adjustment and punishment are $\delta = 1$ and $\lambda = 10$, respectively. Meanwhile, we set the penalty factor for calculating the second-hand reputation score as $\alpha = 0.6$, and the weight for reputation integration as $\sigma = 0.75$. The alarm reputation value for malicious node identification is $R_a = 20$.

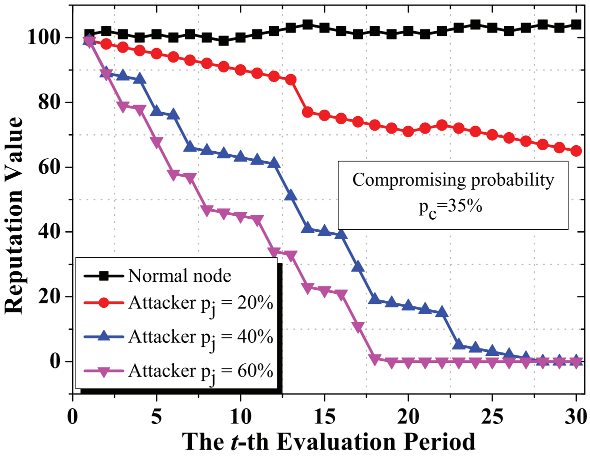

Fig. 5. Reputation value comparison.

A. Reputation Evaluation and Threshold Optimization

CRS-A updates the reputation values of sensor nodes based on their behaviors in data forwarding. The sensor nodes with low reputation values will be identified as malicious nodes over a number of evaluation periods. In Fig. 5, we compare the reputation values of different sensor nodes in 30 evaluation periods. The compromising probability is $P_M = 35\%$ in the simulation. It means that a sensor node has a probability of 35% to be compromised as a malicious node. A larger compromising probability means a larger number of malicious nodes in the network. As shown in the figure, the reputation value of the normal sensor node is slightly increased after 30 periods, while the reputation values of three malicious nodes decrease with different rates. As long as a malicious node increases its attack probability, its reputation value would suffer a dramatic drop after several evaluation periods.

In order to optimize the detection accuracy of CRS-A, we have determined the optimal evaluation threshold in Section V. In Fig. 6, we show the false and missed detection probabilities under different evaluation thresholds, and the determined optimal evaluation period. The simulation results are generated during the transmission from N_j to N_i during an evaluation period, where $S_i(t) = 50$, $p_{i,j}(t) = 30\%$, $P_M = 30\%$ and the attack probability of N_j is $p_j = 40\%$. We can see that with the increment of detection threshold, the missed detection probability $\eta_{i,j}(t)$ increases dramatically but the false detection probability $\mu_{i,j}(t)$ quickly decreases to 0. The optimal threshold is 21 which can make the objective optimizing probability, i.e., $nu_{i,j}(t)$, minimized as 3.1%.

B. Attack Detection Accuracy

In this subsection, we aim to evaluate the detection accuracy of CRS-A by comparing CRS-A with the CAD algorithm [21]. The parameter settings of CAD are adopted according to [21].

[2] Although sensor nodes should send their reputation tables to the sink for malicious node identification, we consider that this part of overhead can be ignored because the interval of malicious node identification is much longer than an evaluation period.

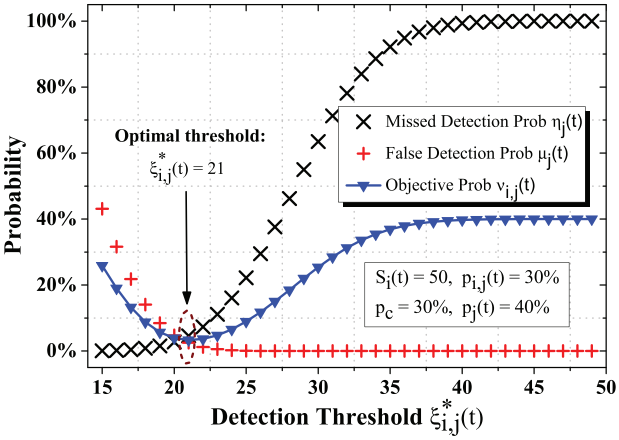

Fig. 6. Evaluation threshold vs False/Missed detection probability.

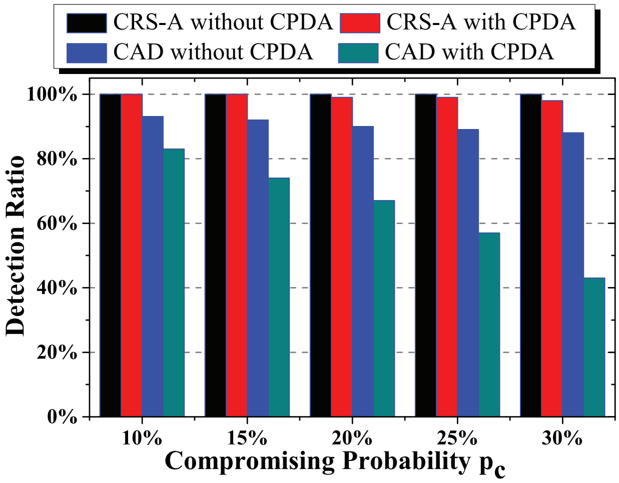

Fig. 7. Detection ratio comparison (CBA means collaborative promotion/demotion attack).

Fig. 7 shows the comparison of malicious node detection ratio in CRS-A and CAD. For both CRS-A and CAD, we compare them in two attack scenarios. The one is that malicious nodes can launch collaborative promotion and demotion attacks (CPDA) to protect each other and defame the normal nodes, while the other is without CPDA. As shown in the figure, CRS-A can detect nearly 100% malicious nodes in both scenarios, while the detection ratio of CAD is much lower in the scenario with CPDA than in the scenario without CPDA. Moreover, even in the scenario without CPDA, the detection ratio of CRS-A is better than CAD's. It demonstrates that CRS-A outperforms CAD in terms of the detection ratio of malicious nodes and is also effective to resist CPDA in attack detection.

Fig. 8 shows the false and missed detection probability comparison between CRS-A and CAD in different compromising probabilities. It can be seen that both the false detection probability and missed detection probability of CRS-A are close to 0 with the increase of the compromising probability. Contrastively, the increment of compromising probability brings a significant ascent in both the false detection probability and missed detection probability.

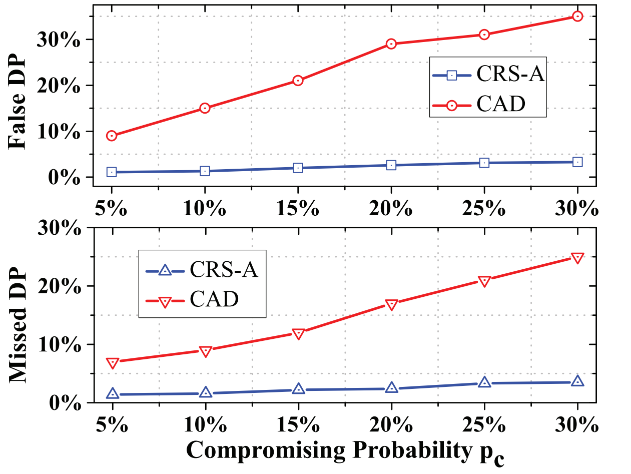

Fig. 8. Detection accuracy comparison. (DP means detection probability in the figure).

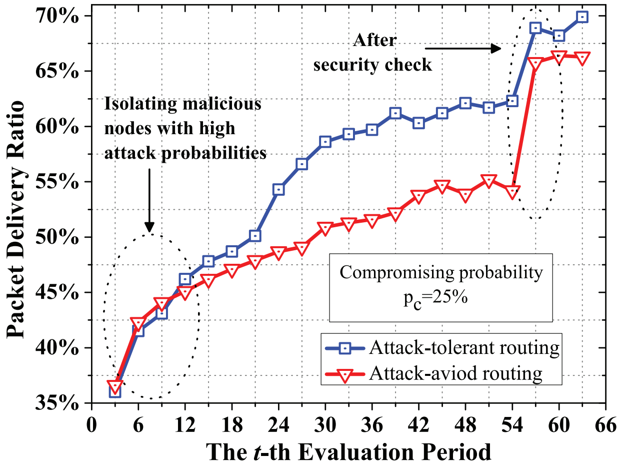

Fig. 9. Packet delivery ratio comparison.

C. Data Delivery Ratio

We evaluate the proposed attack-tolerant routing scheme in this subsection, in terms of the data delivery ratio of the network. In order to show the improvement clearly, we compare the attack-tolerant routing scheme with the attack-avoid routing scheme, where sensor nodes indiscriminately isolate the malicious nodes with low reputation values (below the alarm value R_a) from the routing path. Fig. 9 shows the packet delivery ratio comparison of the two routing scheme over 66 evaluation periods. Both of the routing schemes are applied with CRS-A to the network. The compromising probability of sensor nodes is $P_M = 25\%$, and the security check is performed at the 55-th period, which can update all the identified malicious nodes to be normal. As shown in the figure, the first significant improvement of the data delivery ratio is in the evaluation periods [0, 15] for both routing schemes, where the malicious nodes with high attack probabilities are detected and isolated from the routing path. Meanwhile, after the security check, both routing schemes experience an improvement on data delivery ratio due to the removal of malicious nodes. However, in the periods from 20 to 54, the attack-tolerant routing scheme has a more than 10% improvement on data delivery ratio, compared with

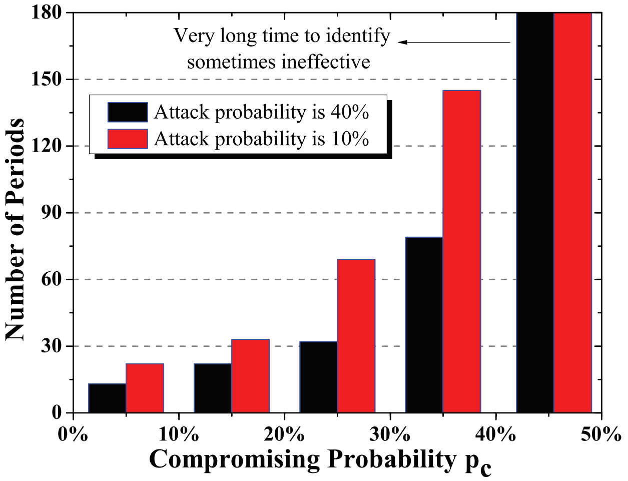

Fig. 10. CRS-A performance for different compromising probabilities.

the attack-avoid routing scheme. That is because the malicious nodes with good channel condition and low attack probabilities are selected into the routing path, and stimulated to perform better to avoid a reputation punishment.

D. Impacts of System Parameters

In this subsection, we evaluate the impacts of system parameters, including compromising probability and R_a, on the performance of CRS-A. Fig. 10 shows the CRS-A performance for different compromising probabilities. To show the detection performance, all the compromised nodes in the simulation are irrational to launch attacks but still can collaborative to protect each other. We aim to compare the number of evaluation periods, within which 90% compromised nodes are identified under different compromising probabilities. The performance is compared under two attack scenarios, where the attacking probabilities of compromised nodes follow two normal distributions with mean values 10% and 40%, respectively. It can be seen that CRS-A can identify the compromised nodes within a very small number of evaluation periods under the compromising probability below 35%. With the increasing compromising probability, the number of evaluation periods increases obviously. Especially, when the compromising probability is 45%, CRS-A has to use a very long time to identify 90% compromised nodes (sometimes cannot identify). It indicates that when there are a large number of compromised nodes, their collaboration can make the performance of CRS-A low and also can make CRS-A ineffective. In addition, it can be observed from the figure that more time should be spent to identify the compromised nodes with low attack probabilities by CRS-A. However, the compromised nodes with low attack probabilities have relatively few impacts on network performance. Fig. 11 shows the impacts of R_a on the performance of CRS-A, in terms of the false identification probability and identification speed. It can be seen that the number of evaluation periods for malicious node identification decreases with the increasing R_a, while the false identification probability increases with the increasing R_a. If WSN applications require the false identification probability below 1%, R_a can be set as 40 to accelerate the malicious node identification while meeting the application requirement.

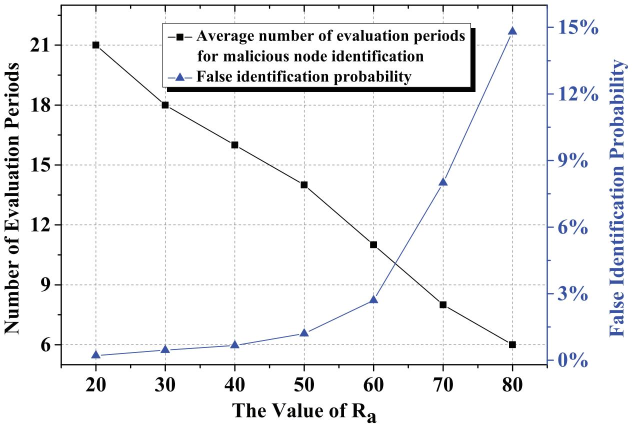

Fig. 11. The Impacts of R_a on the Performance of CRS-A.

VIII. Conclusion

In this paper, we have proposed a channel-aware reputation system with adaptive detection threshold (CRS-A) to detect selective forwarding attacks in WSNs. To accurately distinguish selective forwarding attacks from the normal packet loss, CRS-A evaluates the forwarding behaviors by the deviation between the estimated normal packet loss and monitored packet loss. To improve the detection accuracy of CRS-A, we have further derived the optimal evaluation threshold of CRS-A in a probabilistic way, which is adaptive to the time-varied channel condition and the attack probabilities of compromised nodes. In addition, a distributed and attack-tolerant data forwarding scheme is developed to collaborate with CRS-A for stimulating the cooperation of compromised nodes and improving the data delivery ratio. Our simulation results show that the proposed CRS-A can achieve a high detection accuracy with low false and missed detection probabilities, and the proposed attack-tolerant data forwarding scheme can improve more than 10% data delivery ratio for the network. In our future work, we will extend our investigation into WSNs with mobile sensor nodes, where the detection of selective forwarding attacks becomes more challenging, since the normal packet loss rate is more fluctuant and difficult to estimate due to the mobility of sensor nodes.

References

[1] I. Butun, S. Morgera, and R. Sankar, “A survey of intrusion detection systems in wireless sensor networks,” *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 266–282, May 2014.

[2] Y. Zou, X. Wang, and W. Shen, “Physical-layer security with multiuser scheduling in cognitive radio networks,” *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 5103–5113, Dec. 2013.

[3] B. Xiao, B. Yu, and C. Gao, “CHEMAS: Identify suspect nodes in selective forwarding attacks,” *J. Parallel Distrib. Comput.*, vol. 67, no. 11, pp. 1218–1230, 2007.

[4] Y. Zhang, L. Lazos, and W. Kozma, “AMD: Audit-based misbehavior detection in wireless ad hoc networks,” *IEEE Trans. Mobile Comput.*, Sep. 2013, doi: 10.1109/TMC.2012.257, to be published.

[5] S. Ozdemir, “Functional reputation based reliable data aggregation and transmission for wireless sensor networks,” *Comput. Commun.*, vol. 31, no. 17, pp. 3941–3953, 2008.

[6] D. Hao, X. Liao, A. Adhikari, K. Sakurai, and M. Yokoo, “A repeated game approach for analyzing the collusion on selective forwarding in multihop wireless networks,” *Comput. Commun.*, vol. 35, no. 17, pp. 2125–2137, 2012.

[7] X. Liang, X. Lin, and X. Shen, "Enabling trustworthy service evaluation in service-oriented mobile social networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 2, pp. 310–320, Feb. 2014.

[8] J. Ren, Y. Zhang, K. Zhang, and X. Shen, "SACRM: Social aware crowdsourcing with reputation management in mobile sensing," *Comput. Commun.*, vol. 65, no. 15, pp. 55–65, 2015.

[9] L. Yu, S. Wang, K. Lai, and Y. Nakamori, "Time series forecasting with multiple candidate models: Selecting or combining," *J. Syst. Sci. Complexity*, vol. 18, no. 1, pp. 1–18, 2005.

[10] J. Ren, Y. Zhang, K. Zhang, and X. Shen, "Exploiting channel-aware reputation system against selective forwarding attacks in WSNs," in *Proc. IEEE GLOBECOM*, 2014, pp. 330–335.

[11] S. Djahel, F. Nait-Abdesselam, and Z. Zhang, "Mitigating packet dropping problem in mobile ad hoc networks: Proposals and challenges," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 4, pp. 658–672, Nov. 2011.

[12] K. Liu, J. Deng, P. Varshney, and K. Balakrishnan, "An acknowledgment-based approach for the detection of routing misbehavior in MANETs," *IEEE Trans. Mobile Comput.*, vol. 6, no. 5, pp. 536–550, May 2007.

[13] E. Mahmoud and X. Shen, "An integrated stimulation and punishment mechanism for thwarting packet dropping attack in multihop wireless networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 8, pp. 3947–3962, Oct. 2011.

[14] E. Shakshuki, N. Kang, and T. Sheltami, "EAACK—A secure intrusion-detection system for MANETs," *IEEE Trans. Ind. Electron.*, vol. 60, no. 3, pp. 1089–1098, Mar. 2013.

[15] T. Shu and M. Krunz, "Detection of malicious packet dropping in wireless ad hoc networks based on privacy-preserving public auditing," in *Proc. 5th ACM Conf. Security Privacy Wireless Mobile Netw. (WiSec)*, 2012, pp. 87–98.

[16] S. Marti, T. J. Giuli, K. Lai, and M. Baker, "Mitigating routing misbehavior in mobile ad hoc networks," in *Proc. ACM MobiCom*, 2000, pp. 255–265.

[17] X. Li, R. Lu, X. Liang, and X. Shen, "Side channel monitoring: Packet drop attack detection in wireless ad hoc networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2011, pp. 1–5.

[18] T. Shu, M. Krunz, and S. Liu, "Secure data collection in wireless sensor networks using randomized dispersive routes," *IEEE Trans. Mobile Comput.*, vol. 9, no. 7, pp. 941–954, Jul. 2010.

[19] A. Liu, Z. Zheng, C. Zhang, Z. Chen, and X. Shen, "Secure and energy-efficient disjoint multipath routing for WSNs," *IEEE Trans. Veh. Technol.*, vol. 61, no. 7, pp. 3255–3265, Sep. 2012.

[20] S. Li, S. Zhao, X. Wang, K. Zhang, and L. Li, "Adaptive and secure load-balancing routing protocol for service-oriented wireless sensor networks," *IEEE Syst. J.*, vol. 8, no. 3, pp. 858–867, Sep. 2014.

[21] D. M. Shila, Y. Cheng, and T. Anjali, "Mitigating selective forwarding attacks with a channel-aware approach in WMNs," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1661–1675, May 2010.

[22] Q. Liu, J. Yin, V. Leung, and Z. Cai, "FADE: Forwarding assessment based detection of collaborative grey hole attacks in WMNs," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 5124–5137, Oct. 2013.

[23] J. Ren, Y. Zhang, K. Zhang, and X. Shen, "Exploiting mobile crowdsourcing for pervasive cloud services: Challenges and solutions," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 98–105, Mar. 2015.

[24] A. Nadeem and M. P. Howarth, "A survey of MANET intrusion detection & prevention approaches for network layer attacks," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 2027–2045, Mar. 2013.

[25] H. Lin, X. Zhu, Y. Fang, D. Xing, C. Zhang, and Z. Cao, "Efficient trust based information sharing schemes over distributed collaborative networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 279–290, Sep. 2013.

[26] J. Tang, Y. Cheng, and W. Zhuang, "Real-time misbehavior detection in IEEE 802.11-based wireless networks: An analytical approach," *IEEE Trans. Mobile Comput.*, vol. 13, no. 1, pp. 146–158, Jan. 2014.

[27] N. Baccour *et al.*, "Radio link quality estimation in wireless sensor networks: A survey," *ACM Trans. Sens. Netw.*, vol. 8, no. 4, pp. 1–34, 2012.

[28] T. Liu and A. E. Cerpa, "Data-driven link quality prediction using link features," *ACM Trans. Sensor Netw. (TOSN)*, vol. 10, no. 2, p. 37, 2014.

[29] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[30] IEEE Standard for Information Technology—Telecommunications and Information Exchange between Systems—Local and Metropolitan Area Networks—Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE Standard 802, 2010, pp. 1–51.

[31] K. Govindan and P. Mohapatra, "Trust computations and trust dynamics in mobile ad hoc networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 2, pp. 279–298, May 2012.

[32] R. Shaikh, H. Jameel, B. d'Auriol, H. Lee, S. Lee, and Y.-J. Song, "Group-based trust management scheme for clustered wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 20, no. 11, pp. 1698–1712, Nov. 2009.

[33] J. Ren, Y. Zhang, and K. Liu, "An energy-efficient cyclic diversionary routing strategy against global eavesdroppers in wireless sensor networks," *Int. J. Distrib. Sensor Netw.*, vol. 2013, pp. 1–15, 2013.

[34] J. Ren, Y. Zhang, K. Zhang, A. Liu, J. Chen, and X. Shen, "Lifetime and energy hole evolution analysis in data-gathering wireless sensor networks," *IEEE Trans. Ind. Inf.*, doi: 10.1109/TII.2015.2411231, to be published.

[35] A. Varga and R. Hornig, "An overview of the OMNeT++ simulation environment," in *Proc. 1st Int. Conf. Simul. Tools Techn. Commun. Netw. Syst. Workshop*, 2008, p. 60.

Ju Ren (S'13) received the B.Sc. and M.Sc. degrees in computer science from Central South University, Changsha, China, in 2009 and 2012, respectively. He is currently pursuing the Ph.D. degree in computer science at Central South University. From August 2013 to September 2015, he was also a visiting Ph.D. student in electrical and computer engineering at University of Waterloo, Waterloo, ON, Canada. His research interests include wireless sensor network, mobile sensing/computing, and cloud computing.

Yaoxue Zhang received the B.S. degree from Northwest Institute of Telecommunication Engineering, Xi'an, China, in 1982, and the Ph.D. degree in computer networking from Tohoku University, Sendai, Japan, in 1989. Currently, he is a Professor with the Department of Computer Science, Central South University, Changsha, China, and also a Professor with the Department of Computer Science and Technology, Tsinghua University, Beijing, China. He has authored over 200 technical papers in international journals and conferences, as well as 9 monographs and textbooks. His research interests include computer networking, operating systems, ubiquitous/pervasive computing, transparent computing, and big data. He is a fellow of the Chinese Academy of Engineering and the president of Central South University, China.

Kuan Zhang (S'13) received the B.Sc. degree in electrical and computer engineering and the M.Sc. degree in computer science from Northeastern University, Shenyang, China, in 2009 and 2011, respectively. He is currently pursuing the Ph.D. degree in electrical and computer engineering at University of Waterloo, Waterloo, ON, Canada. His research interests include packet forwarding, and security and privacy for mobile social networks.

Xuemin (Sherman) Shen (M'97–SM'02–F'09) received the B.Sc. degree from Dalian Maritime University, Dalian, China, in 1982, and the M.Sc. and Ph.D. degrees from Rutgers University, New Brunswick, NJ, USA, in 1987 and 1990, respectively, all in electrical engineering. He is a Professor and University Research Chair with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. He is also the Associate Chair for Graduate Studies. His research interests include resource management in interconnected wireless/wired networks, wireless network security, social networks, smart grid, and vehicular ad hoc and sensor networks. He is an elected member of IEEE ComSoc Board of Governor, and the Chair of Distinguished Lecturers Selection Committee. He served as the technical program committee Chair/Co-Chair for the IEEE Globecom'16, the Infocom'14, the IEEE VTC'10 Fall, and the Globecom'07, the Symposia Chair for the IEEE ICC'10, the Tutorial Chair for the IEEE VTC'11 Spring and the IEEE ICC'08, the General Co-Chair for ACM Mobihoc'15, Chinacom'07 and QShine'06, the Chair for the IEEE Communications Society Technical Committee on Wireless Communications, and P2P Communications and Networking. He also serves/served as the Editor-in-Chief for the *IEEE Network, Peer-to-Peer Networking and Application*, and *IET Communications*; a Founding Area Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS; an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, *Computer Networks*, and *ACM/Wireless Networks*, etc.; and the Guest Editor for the IEEE JSAC, the IEEE WIRELESS COMMUNICATIONS, the *IEEE Communications Magazine*, and *ACM Mobile Networks and Applications*, etc. He was the recipient of the Excellent Graduate Supervision Award in 2006, and the Outstanding Performance Award in 2004, 2007, 2010, and 2014 from the University of Waterloo, the Premier's Research Excellence Award (PREA) in 2003 from the Province of Ontario, Canada, and the Distinguished Performance Award in 2002 and 2007 from the Faculty of Engineering, University of Waterloo. He is a registered Professional Engineer of Ontario, Canada, an Engineering Institute of Canada Fellow, a Canadian Academy of Engineering Fellow, a Royal Society of Canada Fellow, and a Distinguished Lecturer of the IEEE Vehicular Technology Society and Communications Society.