CHANNEL AWARE DETECTION OF FORWARDING ATTACKS IN WSN WITH MALICIOUS NODE DETECTION BASED ON CO-OPERATIVE APPROACH

Ragitha R

M.tech student, Dept. of Electronics & Communication Engineering, KiTS college of engineering, APJ Abdul Kalam technological University,

---

Abstract: *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 battle field reconnaissance and homeland security monitoring. So due to the lack of physical protection, sensor nodes are easily compromised by adversaries. Wireless sensor networks (WSNs) are vulnerable to selective forwarding attacks that can maliciously drop a subset of forwarding packets to degrade network performance. 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. A Channel-aware Reputation System with adaptive detection threshold (CRS-A) can 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. 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. But these solutions for wireless networks may not always be sufficient. Because some malicious nodes pretend to be intermediate nodes of a route to some given destinations, drop any packet that subsequently goes through it, is one of the major types of attack. In this paper, in addition to CRS-A, uses Ad-hoc on demand Distance Vector (AODV) routing that propose a co-operative method to detect malicious node effectively.*

Key words: WSN, AODV, CRS-A, malicious node, reputation value, co operative approach

1. 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 battle field 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. 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 challenging to detect the selective forwarding attacks and improve the network performance.

In this paper, I propose a Channel-aware Reputation System with adaptive detection threshold (CRS-A) [1] to detect selective forwarding attacks in WSNs with the detection of malicious node. Specifically, we divide the network lifetime to a sequence of evaluation periods. 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. In the malicious node detection phase, each node transmits data to a next node, stores a copy of the data in its buffer and overhears whether the next node transmits the data. If the node overhears data transmission of the next node within a predetermined length of time, the node considers that the data was properly transmitted and deletes the copy of the data from the buffer. If not so, the node increases a failure tally for the next node. If the failure tally is greater than a threshold, the node determines that the next node intentionally dropped the data and reports this fact to all nodes over the network. The mechanism is cooperative because nodes in the protocol work co-operatively together so that they can analyze, detect malicious nodes in a reliable manner.

2. EXISTING TECHNIQUES

2.1 ADAPTIVE AND CHANNEL AWARE DETECTION OF SELECTIVE FORWARDING ATTACKS IN WSN

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.

2.2 DETECTING MALICIOUS NODES IN MANET BASED ON A COOPERATIVE APPROACH

Mobile Ad hoc Network (MANET) is a self-configuring network of mobile nodes connected by wireless links and considered as network without infrastructure. Securing MANETs is an important part of deploying and utilizing them, since they are often used in critical applications where data and communications integrity is important. Existing solutions for wireless networks can be used to obtain a certain level of such security. These solutions may not always be sufficient, as ad-hoc networks have their own vulnerabilities that cannot be addressed by these solutions. In the network, some malicious nodes pretend to be intermediate nodes of a route to some given destinations, drop any packet that subsequently goes through, it is one of the major types of attack. We propose a cooperative method to detect malicious nodes in MANETs. The mechanism is cooperative because nodes in the protocol work cooperatively together so that they can analyze, detect malicious nodes in a reliable manner. We verify our method by running simulations with mobile nodes using Ad-hoc on demand Distance Vector (AODV) routing. It is observed that the malicious node detection rate is very good; the overhead detection rate is low, packet delivery ratio is little bit high and also the response time is observed when there is a change of mobility speed. In this, each node transmits data to a next node, stores a copy of the data in its buffer and overhears whether the next node transmits the data. If the node overhears data transmission of the next node within a predetermined length of time, the node considers that the data was properly transmitted and deletes the copy of the data from the buffer. If not so, the node increases a failure tally for the next no de. If the failure tally is greater than a threshold, the node determines that the next node intentionally dropped the data and reports this fact to all nodes over the network. Each of the nodes receiving the report determines whether a reporter and a suspect node listed in the report are recorded in its report table. When the number of times that a node reports to the source node S is greater than k equivalent to the number of malicious nodes over the network, the node is determined as a malicious node and excluded from the network.

Most of these systems 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. 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.

3. SYSTEM MODEL AND DESIGN GOALS

We consider a WSN consisting of a set of randomly distributed sensor nodes, denoted by 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. 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 PM 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 or acknowledgment- based approaches. It means that a sensor node can obtain that how many data packets are forwarded by its forwarding sensor nodes. Existing works 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.

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.

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

We consider that cryptographic techniques have been utilized in the network to provide sufficient data confidentiality and authentication against the adversary, and 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. 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. 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.

4. CRS-A: THE CHANNEL-AWARE REPUTATION SYSTEM WITH ADAPTIVE DETECTION THRESHOLD

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 T = {T1,...,Tt,...}. In each evaluation period Tt, the channel condition of each data transmission link is assumed to be stable. Meanwhile, for each Tt, we introduce a channel estimation stage at the beginning of Tt, and a reputation update stage at the end of Tt.

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. 4.1 shows the overview of evaluation periods over the network lifetime. 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 them and the propagated reputation scores from their neighboring nodes to update the reputation table.

4.1 Normal Packet Loss Estimation

According to the network model, normal packet loss is mainly caused by the poor and unstable wireless channel and MAC layer collisions. The poor and unstable radio link quality is the primary reason for the time-varied packet losses. It 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. 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. 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.

As data transmission between two neighboring nodes is based on the IEEE 802.11, 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 to estimate the packet loss caused by medium access collisions without the impact of hidden terminals. 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

$$p_i = (1-p_t)^n$$

$$p_s = n.p_t.(1-pt)^{n-1}$$

$$p_c = 1-p_i-p_s$$

And the channel busy ratio Rb can be calculated as

$$Cb = 1-(p_i \,.\, t_d)/(p_i \cdot \sigma + p_s \,.\, t_s + p_c \,.\, t_c)$$

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.

4.2 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 is named as first-hand reputation scores. Specifically, in the data transmission stage of T_t, node N_i ($N_i \in N$) 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.

$$P_{i,j}(X = k) = ({}^{l}{}_{k})(p_{i,j}(t))^k(1- p_{i,j}(t))^{l-k.}$$

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 $\xi_{i,j}(t)$ ($S_{i,j}(t)\cdot p_{i,j}(t) < \xi_{i,j}(t) < S_{i,j}(t)$, $\xi_{i,j}(t) \in N+$) and define the reputation evaluation function of N_i to N_j as follows.

$$r^1{}_{i,j}(t) = \begin{cases} +\delta, & if\ mi,j(t) \le pi,j(t)\cdot Si,j(t) \\ -\delta, & if\ pi,j(t)\cdot Si,j(t) < mi,j(t) \le \xi i,j(t) \\ -\lambda & if\ mi,j(t) > \xi i,j(t) \end{cases}$$

where λ is a punishment factor and δ is a adjustment factor. We set and explain the function as follows.

- If $mi,j(t) \le p_{i,j}(t)\cdot S_{i,j}(t)$, 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 to N_j to neutralize the reputation evaluation.

- When $mi,j(t) > \xi_{i,j}(t)$ we consider there is a high probability for N_j to misbehave in the data forwarding. If it happens, Ni rates a punishment - λ to N_j.

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.

4.3 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. 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 $NC_{i,b}$. 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. 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.

4.4 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.

$$Ri,j = \begin{cases} Rs, & \text{if } Ri,j + RIi,j \leq Rs \\ Ri,j + RIi,j, & \text{if } Rs < Ri,j + RIi,j < Rm \\ Rm, & \text{otherwise} \end{cases}$$

Here, σ is the weight factor of the first-hand information and σ>0.5. R_m and R_s are system parameters that can be chosen based on the system requirements.

4.5 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, N_j is identified as a malicious node. Here, Ra 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.

5. ADAPTIVE DETECTION THRESHOLD FOR CRS-A

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 indicate 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.

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. 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 . 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 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.

Here 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

$$\eta_{i,j}(t) = P\{X+Y \leq \xi_{i,j}(t) / \text{ j misbehaved in } T_t\}$$

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

$$\mu_{i,j}(t) = P\{X+Y \leq \xi_{i,j}(t) / \text{ j behaved well in } T_t\}$$

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 μi,j(t). X is defined as the number of normally lost data packets during the transmission. 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.

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 PM 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,($0 \le a \le S_{i,j}(t)$, $0 \le k \le a$) as

$$P\{Y' = k \mid A_i(t)=a\}= \binom{a}{k} p_j^k (1-p_j)^{a-k}$$

And the PMF of Y' is

$$P\{Y' = k\}= \sum_{a=0}^{Si,j(t)} [P\{Y' = k \mid A_i(t)=a\}\ P\{A_i(t)=a\}]$$

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

$$P\{Y = k\} = \begin{cases} (1-PM) + PM.P\{Y' = 0\}, if\ k = 0 \\ PM.P\{Y' = k\}, \quad if\ 1 \le k \le Si,j(t) \end{cases}$$

If Ni sends $S_{i,j}(t)$ data packets to Nj during the evaluation period Tt and the detection threshold is $\xi i,j(t)$ ($Si,j(t)\cdot pi,j(t) < \xi i,j(t) < Si,j(t)$), the missed detection probability for evaluating Nj is

$$\eta_{i,j}(t) = \frac{\sum_{k=1}^{\xi i,j(t)} [P\{X \le \xi i,j(t)-k\} . P\{Y=k\}, \text{if } k=0}{P_M - P_M . (1-p_j)^{Si,j(t)}}$$

Where $P\{X \le k\}$ is the CDF of X

Based on the PMF of X and Y, we further calculate η_j as follows.

$$\eta_j = \frac{P\{\{X+Y \le \xi i,j(t)\} \cap \{Y > 0\}\}}{P\{Y > 0\}}$$

$$= \frac{P\{\{X+Y \le \xi i,j(t)\} \cap \{\sum_{k=1}^{Si,j(t)} \{Y = k\}}{P\{\sum_{k=1}^{Si,j(t)} \{Y = k\}}$$

$$= \frac{\sum_{k=1}^{Si,j(t)} P\{\{X+Y \le \xi i,j(t)\} \cap \{Y=k\}}{P\{\sum_{k=1}^{Si,j(t)} \{Y = k\}}$$

$$= \frac{\sum_{k=1}^{Si,j(t)} [P\{X+Y \le \xi i,j(t) \mid Y = k\} . P\{Y= k\}]}{P\{\sum_{k=1}^{Si,j(t)} \{Y = k\}}$$

The missed detection probability $\eta_{i,j}(t)$ 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

$$p_j = \frac{[\sum_{w=0}^{t} [mi,j(w) - Si,j(w).(1-pi,j(w))]]}{\sum_{w=0}^{t} [Si,j(w).(1-pi,j(w))]}$$

Where $S_{i,j}(w).(1-p_{i,j}(w))$ is the expected number of forwarded data packets at time period w, while $m_{i,j}(w) - S_{i,j}(w)\cdot(1-(p_{i,j}(w))$ is deviation between the actual number of forwarded data packets and the expected number of forwarded data packets at time period w. 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 $\eta_{i,j}(t)$ should be emphasized to optimize the performance of CRS-A. Meanwhile, both of the missed detection probability $\eta_{i,j}(t)$ and false detection probability μ_j depend on $\xi_{i,j}(t)$. When $\xi_{i,j}(t)$ increases, $\eta_{i,j}(t)$ increases and μ_j decreases. And if $\xi_{i,j}(t)$ decreases, the situation reverses. It means $\eta_{i,j}(t)$ and μ_j are two contradictory optimization objectives. In order to find a trade-off between them, we can integrate $\eta_{i,j}(t)$ 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 $\nu_j = p_j \cdot \eta_{i,j}(t) + (1-p_j)\cdot\mu_j$. 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

(PP) minimize $\nu_j = p_j.\eta_{i,j}(t) + (1-p_j)\cdot\mu_{i,j}(t)$

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(Si(t)), i.e., O(n).

6. 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 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. 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.

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. In this 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 Figure 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, Ns have 6 data packets to forward. If Ns choose Nb as the next hop, the expected number of data packets that are successfully forwarded by Nb is 3. Contrastively, the expected number of data packets forwarded by Na should be 5, even if its reputation in Ns is low and it has an attack probability 20% according to the historical records.

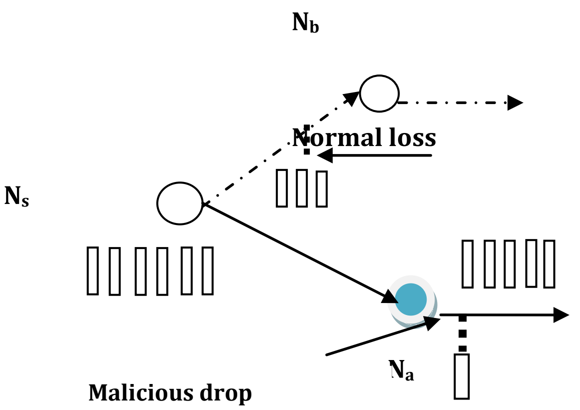

Figure 3 Example of dynamic routing

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

$$DFR_j(t) = (S_i(t)L_j(t))/S_i(t)$$

$$= 1 - p_{i,j}(t) - p_j(t) + p_{i,j}(t)\cdot p_j(t)$$

4.5.1 Algorithm

Description: Updating the reputation of sensor nodes and data forwarding during T_t ($T_t \in T$).

Phase I Normal Loss Estimation;

for each Ni∈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 Ni ∈N **do**

Choosing N_j from RC_i as the next hop 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 Ni ∈N **do**

Calculate the attack probability p_j of N_j;

Determine the optimal detection threshold $\xi_{i,j}(t)$ by solving the problem (PP);

Evaluate the first-hand reputation score $r^1{}_{i,j}(t)$

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)

end

Calculate the integrated reputation score $R^I{}_{i,j}(t)$ with $r^1{}_{i,j}(t)$ and $r^2{}_{i,j}(t)$ and update $R_{i,j}$

end

According to Algorithm 1, when a malicious node N_j is selected into the routing path by Ni, 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.

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 Ni for maintaining the CRS-A is $8 \cdot |NC_i|$ bits, where NC_i is the neighbor set of N_i. The communication overhead is mainly produced by channel estimation and reputation propagation. Let B be the number of bits in a PROBE packet that sensor nodes broadcast to their neighboring nodes for channel estimation. 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.

7. MALICIOUS NODE DETECTION

To detect the malicious node we have proposed one method which uses a reactive routing protocol known as Ad hoc On demand Distance Vector (AODV) routing for analysis of the effect of the black hole attack when the destination sequence number is changed via simulation. The proposed algorithm first detects those nodes, which may be malicious. Then the neighbor of the malicious node initiates a cooperative detection mechanism to detect the actual black hole node. In AODV routing, messages contain only the source and the destination addresses. It uses destination sequence numbers to specify the valid route. At first the sender broadcast the Route Request (RREQ) message to its neighbors. Each node that receives the broadcast, checks the destination to see if it is the intended recipient. If yes it sends a Route Reply (RREP) message back to the originator. RREP message contains the current sequence number of the destination node. The same process continues till the packets reach to destination or reach to an intermediate node, which has a fresh, enough routes to destination. Every node keeps track of its neighbor by maintaining two small size tables. One is sequence table (SnT) to keep the neighbor node's id and neighbor node's sequence number and other is the status table (ST) to keep track of the node's status whether it is a safe node or a malicious one. Every node also maintains a neighbor list (N_List) and this list is updated periodically. When an intermediate node receives a RREP checks if the difference between the Dst_Seq present in the RREP message and the sequence no present in its table is greater than some predefined threshold value? if so then the intermediate node stops forwarding the message and mark the node as „M" or malicious in the status table(ST) and send a notification message(NM) to source node along with the malicious node's id and neighbor list of the malicious node. The threshold value is the average difference of Dst_Seq in each time slot between the sequence number of RREP message and the one held in the table.

The source node has an additional table called Flag Table (FT). M1HN's after receiving the Further Detection message, broadcast a RREQ message by setting destination address to source node's address. If it receives a RREP message from the malicious node, it sends a Test packet (TP) to the source node via malicious node, and at the same time it sends a Acknowledgment Packet (AP) to source node(SN) though some other route. Then the source node waits for *wt* time until it receives the entire test and acknowledgement packet. If, SN receives a TP, it updates the Flag Table (FT) by adding the source node id to the table and set the flag of the node as Y and if an AP is received set the flag as N and update the count field. If all the entries for the malicious node are N then source node updates the status table (ST) by adding the MN s id to the ST and making the status as B i.e. Black hole.

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.

8. CONCLUSION

WSN is being emerged as a promising and interesting area. It is designed for real-time data collection and analysis of data in hostile environments so they are used mainly in monitoring and surveillance based applications. Most widely used applications of WSN are military appliance, area monitoring, environmental monitoring, industrial monitoring, machine health monitoring, water/waste water monitoring, and fleet monitoring. Since, WSNs are mostly used in a hostile environment security is mainly concerned. The conventional security measures are not suitable to the wireless sensor networks due to resource constraints of both memory and energy. In WSN, sensor nodes use wireless communication to send packets. A sensor node uses multi-hop transmission to deliver the packet to the base station, due to its limited transmission range. So a packet is forwarded through too many hops/nodes to reach the destination. As, we discussed sensor networks are usually deployed in hostile environments, an adversary can launch attacks. Attacks can be classified into two types, inside attacks and outside attacks. The latter one can be easily detected and security solutions are provided. In former one, adversary compromises some internal nodes and launches attacks which will be difficult to detect.

This paper proposes two functions; they are the channel aware detection of forwarding attacks using CRS-A and the malicious node detection based on co-operative approach.

Experiments are conducted using NS-2 version 2.35, a scalable simulation environment for network systems. The routing protocol we use is AODV. Our simulated network consists of 50 mobile nodes placed randomly with all nodes have the same transmission range. The channel capacity is 2 Mbps. We setup the parameters of CRS-A as follows. The range of the reputation value of a sensor nodes is [0,200], i.e., Rs =0 and Rm = 200. The initial reputation is 100 for all the sensor nodes. The value of adjustment and punishment are δ=1and λ= 10, respectively. Meanwhile, we set the penalty factor for calculating the second-hand reputation score as α =0 .6, and the weight for reputation integration as σ =0 .75. The alarm reputation value for malicious node identification is Ra = 20.

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. The compromising probability is PM = 40% in the simulation. It means that a sensor node has a probability of 40% to be compromised as a malicious node. A larger compromising probability means a larger number of malicious nodes in the network.

REFERENCES

[1] “Adaptive and Channel-Aware Detection of Selective Forwarding Attacks in Wireless Sensor Networks” by Ju Ren, Student Member, IEEE, Yaoxue Zhang, Kuan Zhang, Student Member, IEEE, and Xuemin (Sherman) Shen, Fellow, IEEE

[2] “Detecting Malicious Nodes in MANET based on a Cooperative Approach” by Reena Sahoo Dept. Of Information Technology Sambalpur University Sambalpur, Odisha, India and Dept. Of Information Technology Sambalpur University Sambalpur, Odisha, India and Dr. P. M. Khilar Dept. of CSE NIT, Rourkela, Odisha, I

[3] B. Yu and B. Xiao, “Detecting selective forwarding attacks in wireless sensor networks,” in Proc. of the 2nd International Workshop on Security in Systems and Networks, April 2006, pp. 1-8.

[4] Tran Hoang Hai, Eui-Nam Huh, “Detecting Selective Forwarding Attacks in Wireless Sensor Networks Using Two-hops Neighbour Knowledge” Seventh IEEE International Symposium on Network Computing and Applications, 2008, pp.325-331.