Throughput Analysis and Routing Security Discussions of Mobile Access Coordinated Wireless Sensor Networks

Mai Abdelhakim Jian Ren Tongtong Li
Department of Electrical & Computer Engineering, Michigan State University, East Lansing, MI 48824, USA.
Email: {abdelhak, renjian, tongli}@egr.msu.edu

***Abstract*—In this paper, we analyze the throughput of a novel mobile access coordinated wireless sensor network architecture (MC-WSN) under single path and multipath routing. The obtained throughput expressions highlight the trade-off between achieving high throughput performance and improving the network security strength. The results reveal the importance of: (i) minimizing the number of hops in maximizing the throughput, and (ii) adopting routing diversity in combating malicious attacks and network failure conditions. We control the number of hops in data transmission through optimal topology design and active network deployment achieved by the mobile access point (MA). To combat routing attacks, we propose a secure routing path selection approach, and show the impact of the proposed approach on improving the throughput performance under malicious attacks.**

***Index Terms*—Wireless sensor networks, mobile access coordinator, throughput, routing security.**

I. Introduction

Wireless sensor network has received significant research attentions due to their paramount impact on various critical military and civilian applications. The adopted architecture largely impacts the performance of the network. For example, in random ad-hoc networks, it was shown in [1] that the average throughput of a node vanishes as the number of nodes in the network increases. This indicates that for efficient communications, the network should not be completely structureless. In [2], throughput of multihop many-to-one network with uniformly deployed nodes and time division multiple access (TDMA) protocol is investigated, and it was shown that clustering the network can potentially improve the throughput.

Incorporating mobility in the network architecture could improve the throughput performance. For example, when the network nodes are mobile and the packets are split over multiple mobile relays, the achieved throughput can be higher than that of fixed ad-hoc networks [3]. The drawback, though, is higher energy consumption, which is a big challenge for sensor networks. When the sinks or access points are mobile, like in Sensor Networks with Mobile Access Points (SENMA) [4], the capacity of a node having a direct communication with the sink is significantly superior to that of ad-hoc sensor networks. However, since the sources wait for an access point's visit, the throughput is limited by the traversal speed and trajectory length of the access point.

In this paper, we consider the novel Mobile Access Coordinated Wireless Sensor Network Architecture (MC-WSN). In MC-WSN, multiple sinks are utilized for data collection. There exist a powerful center cluster head (CCH) located in the middle of each cell, and powerful ring cluster heads (RCHs) that are uniformly placed on a circular ring in each cell. Data transmission goes through simple routing from each sensor to the CCH and/or RCHs. MC-WSN also exploits mobile access points (MAs) to coordinate the network by deploying sensors nodes and cluster heads, recharging low-power nodes, as well as detecting malicious sensors and replacing them. MAs collects data from the sensors through the CCH and RCHS. The MA can also traverse the cell if the routing paths do not work. The feasibility of this design is supported by the recent advances in unmanned aerial vehicle technology [5].

In this paper, we extend our previous work in [6], and calculate the throughput of the MC-WSN. We start by analyzing the throughput from an information theoretic perspective, then obtain the throughput expressions under both single path and multiplath routing scenarios. Moreover, we discuss routing security in MC-WSN. Sensor networks are vulnerable to malicious attacks that could severely degrade the network performance. Among the different attacks, the compromisation of authenticated nodes poses serious security threats. This includes routing attacks, where malicious nodes can modify and drop packets that are being routed to the sink [7]. In this paper, we propose a secure routing path selection mechanism to mitigate routing attacks in MC-WSN, and highlight the inherent security features of MC-WSN in combating these attacks. We then discuss the throughput limitation of MC-WSN under routing attacks, and the improvements achieved through employing the proposed secure routing path selection approach.

II. The Mobile Access Coordinated Wireless Sensor Network (MC-WSN)

In this section, we describe the MC-WSN architecture, present its major advantages and features, and discuss the hop number control.

A. Description of the Architecture

In MC-WSN, we assume the network is divided into cells each of radius d. Each cell contains a single powerful mobile

978-1-4799-3512-3/14/$31.00 ©2014 IEEE

access point (MA) and n uniformly deployed sensor nodes (SNs) that are arranged into K_{CH} clusters. Each cluster is managed by a cluster head (CH), to which all the cluster members report their data. CHs then route the data to the MA. A powerful center cluster head (CCH) is employed in the middle of each cell, and K powerful ring cluster heads (RCH) are placed on a ring of radius R_t. The CCH and RCHs can establish direct communication with the MA or with other RCH that are closer to the MA. All nodes within a distance R_o from the CCH route their data to the MA through the CCH. All other nodes route their data to the MA through the nearest RCH. If a sensor is within the MA's coverage range, then direct communications can take place. After receiving the data of the sensors, the MA delivers it to a Base Station (BS). The overall network architecture is illustrated in Figure 1. As will be illustrated later, the number of hops from any sensor to the MA can be limited to a pre-specified number through the deployment of CCH and RCHs.

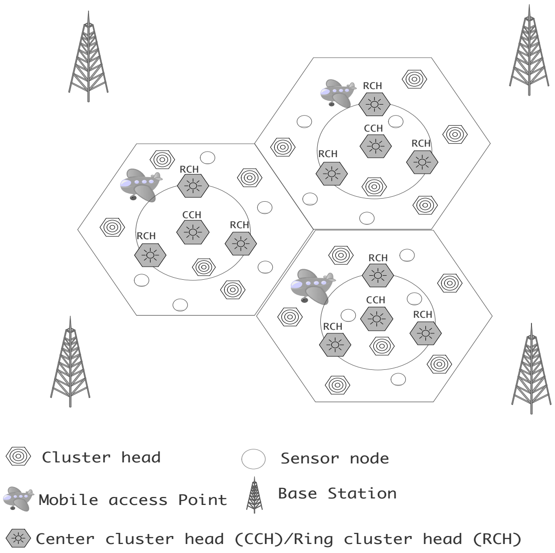

Fig. 1. Proposed MC-WSN architecture.

In the proposed MC-WSN architecture, the MA coordinates the sensor network and resolves the node deployment issue as well as the energy consumption problem of wireless sensor networks. More specifically, the MAs are responsible for (i) deploying nodes, (ii) replacing and recharging nodes, (iii) detecting malicious sensors, then removing and replacing them, (iv) collecting the information from sensors and delivering it to the BS.

The MAs can move on the ground, and can also fly. Each MA traverses its cell mainly for removing the malicious nodes and replacing or recharging low-energy sensor nodes and cluster heads. It moves physically for data collection only in the case when the routing paths do not work.

Data transmission from any SN to the MA goes through simple routing, either with the CCH or the RCHs. Let the communication range of each sensor node and CH be r_c and R_c, respectively. SNs only communicate with their corresponding CH, which then routes their data to the MA. CHs have larger storage capacity and longer communication range than SNs, i.e., $R_c > r_c$.

The main advantages of MC-WSN lie in: the multi-functionality of the mobile access, the hop number control through topology design, and hierarchical and heterogeneous sensor deployment. MC-WSN has the following features:

- Controlled network development and prolonged network lifetime
- Time-sensitive data transmission
- Enhanced network security
- Efficient energy consumption
- Enhanced network resilience, reliability and scalability:

Due to the active network deployment feature in the proposed architecture, we can assume that the nodes are uniformly distributed in the network. It is also reasonable to place the powerful RCHs at evenly spaced locations on the ring R_t. Figure 2 shows an example of the MC-WSN with four RCHs.

Under normal network conditions, i.e., without malicious attacks, to maximize the throughput and minimize the delay of data transmission from the sensors to the MA, the number of hops needed in routing should be minimized. In Section II-B, we consider topology design for hop number minimization.

B. Hop Number Minimization

In this subsection, we obtain the optimal radius R_o and the ring radius R_t that minimize the average number of hops from any CH to its nearest sink. Assuming multiple paths are present due to network diversity, this topology design will ensure that at least one of the paths from any node to a sink has minimum number of hops, i.e., shortest path. Shortest path routing results in higher throughput performance under normal network conditions as will be illustrated in Section III. Note that under shortest path routing, the number of hops is proportional to the distance between the source and the sink. To minimize the number of hops, *we design the topology such that the average distance between a cluster head and its nearest sink is minimized.*

We calculate the average squared distance between any source and the corresponding sink (CCH/RCH) $\bar{d}^2$ according to Fig.2, set $\frac{\partial \bar{d}^2}{\partial R_o} = 0, \quad \frac{\partial \bar{d}^2}{\partial R_t} = 0$, and get the following result:

***Proposition** 1: Assuming a circular cell of radius* d*, to minimize the number of hops in the MC-WSN architecture with one CCH and* K *RCHs, where* $K > 1$*, data transmission should be arranged as follows: (1) The CHs within a distance* $R_o = 0.366\ d$ *from the center of the cell deliver their data to the MA through the CCH. (2) The CHs at a distance* x *from CCH, where* $R_o \leq x < d$*, deliver their data to the MA through the nearest RCH on the ring of radius* $R_t = 0.233K\sin(\frac{\pi}{K})d$.

The average number of hops along the shortest path routing can be expressed as $N_{hop} = \frac{\bar{d}}{R_c}$, where R_c is the communication range of the cluster heads. Note that as K increases, $\bar{d}$ and consequently N_{hop} decrease. On the other hand, if the cell radius d decreases, then also $\bar{d}$ and N_{hop} decrease. As can be seen, the maximum number of hops can be limited to a pre-specified number through active deployment of RCHs.

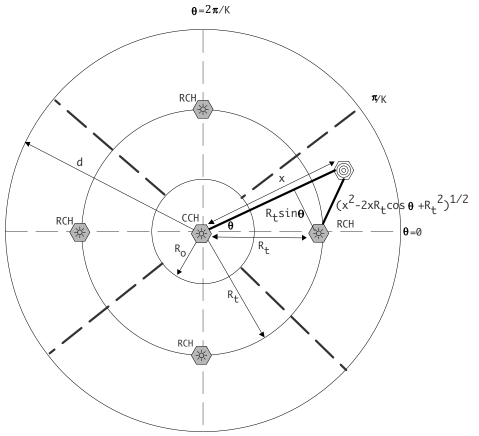

Fig. 2. MC-WSN with four powerful RCHs.

III. Throughput Analysis

In this section, we analyze the throughput of the multihop MC-WSN architecture. We first introduce the definition of the throughput in the single hop case, then analyze the multihop throughput under both single path and multipath routing.

A. Brief Overview

As an important measure of network performance, throughput is generally defined as the amount of information that can be successfully transmitted over a network, and is largely determined by the network model and transmission protocols. Existing work on throughput analysis is versatile [1], [8], [9], including one-hop centralized cases [8], [9] and ad-hoc cases [1]. There are also research on systems with mobile nodes [3] and systems with mobile access points, like SENMA [4]. In SENMA, as there is a direct link between each sensor and the mobile sink, the system throughput is significantly superior to that of ad-hoc sensor networks [4].

In [1], the throughput of random ad-hoc networks is studied. It was shown that the throughput obtained by each node vanishes as the number of nodes in the network increases. More specifically, for an ad-hoc network containing n nodes, the throughput obtainable by each node is $O(\frac{W}{\sqrt{n}})$ bit-meters/sec, where W is the maximum capacity of each link in the network. Note that the size or density of an ad-hoc network or a wireless sensor network plays a critical role in network performance. This result indicates that for reliable and efficient communications, the network cannot be completely structureless, but should have a well-defined structure while maintaining sufficient flexibility. This thought has actually been reflected in the merging of centralized and ad-hoc networks, leading to ad-hoc networks with structures, known as hybrid networks [10]. As can be seen, the proposed MC-WSN is also an example of hybrid network: it has a hierarchical structure supported by the CCH, RCHs, and CHs; at the same time, it also allows partially ad-hoc routing for network flexibility and diversity.

B. Definition of the Throughput

We start with the single hop case. Assuming node i is transmitting to sink k, where $k \in \{0, 1, ..., K\}$. *The throughput of node i to sink k*, $T_{i,k}$, is defined as the average number of packets per slot that are initiated by node i and successfully delivered to the intended receiver k. Define $R_S^k(\nu)$ as the set of nodes that have their packets successfully delivered to sink k in slot ν, where S is the set of nodes scheduled to transmit. Then, $T_{i,k}$ can be expressed as: $T_{i,k} = \lim_{T\to\infty} \frac{1}{T} \sum_{\nu=1}^{T} Pr\{i \in R_S^k(\nu)\}$ [11].

Assuming that the packet reception from slot to slot is an i.i.d process, then the throughput can be formulated using two binary flags t_i^k and r_i^k, where t_i^k is a binary flag indicating that node i transmits data to sink k and r_i^k is a binary flag indicating that the data of node i is successfully received at the intended destination k (CCH or RCH). It follows that:

$$T_{i,k} = Pr\{r_i^k = 1 | t_i^k = 1\} Pr\{t_i^k = 1\}. \tag{1}$$

Note that the transmission from the powerful CCH/RCH to the MA can be made at high-power and high-rate. Also, with the active network deployment performed by the MA, the data from each sensor to its CH can be transmitted over a single hop using a collision-free MAC protocol. Thus, we focus on data transmission from the CH of the originating node to its corresponding CCH/RCH.

In the following, we analyze $T_{i,k}$ from the information theory perspective, by discussing the relationship between $T_{i,k}$ and the mutual information between the packet transmitted from CH i and the packet received at sink k.

For each slot, define X_i^k as the transmitted packet from CH i to sink k, where $X_i^k = 0$ means that node i is not transmitting. Let $\tilde{X}_i^k$ be the non-zero packets of X_i^k, then $X_i^k = t_i^k \tilde{X}_i^k$ [8]. Assuming that sink k receives packets from multiple nodes. Define $\mathbf{Y}^k$ as the received vector at sink k, where the ith element in $\mathbf{Y}^k$ is the received packet from CH i. Let $\mathbf{r}^k$ be the vector whose ith element is r_i^k. It has been shown in [8] that the mutual information between X_i^k and $\mathbf{Y}^k$ can be written as a function of the throughput of CH i to sink k ($T_{i,k}$) as follows:

$$\mathbf{I}(X_i^k, \mathbf{Y}^k) = \mathbf{I}(t_i^k, \mathbf{r}^k) + H(\tilde{X}_i^k) T_{i,k}, \tag{2}$$

where $\mathbf{I}(x, y)$ is the mutual information between x and y, and $H(x)$ is the entropy of x. Let $\mathbf{I}_p^k = \mathbf{I}(X_i^k, \mathbf{Y}^k)/H(\tilde{X}_i^k)$, which is measured in number of packets per slot. In general, $T_{i,k} \leq \mathbf{I}_p^k$. Note that t_i^k is binary, i.e., $H(t_i^k) \leq 1$, which implies that $\mathbf{I}(t_i^k, \mathbf{r}^k) \leq H(t_i^k) \leq 1$. As a result, if the packet length gets large, i.e., $H(\tilde{X}_i^k) \to \infty$, then we have $T_{i,k} \simeq \mathbf{I}_p^k$.

From the information theory perceptive, this shows that: $T_{i,k}$ is the average normalized information (measured in packets per slot) passed through the channel between CH i and sink k.

C. Multihop Single Path Routing Case

Consider that CH i requires N_i^k hops to reach sink k. N_i^k is based on the network architecture, topology, and routing scheme. Let the ideal or shortest path from CH i to sink k be $i_{N_i^k} \to i_{N_i^k - 1} \to ... i_1 \to i_0$, where $i_{N_i^k}$ is the source CH i and i_0 is the sink k. Let $t_{i,h}^k$ be a binary flag at hop h, indicating that CH i_h is scheduled to relay a packet of CH i

to CH i_{h-1} along the route to sink k. Also, let $r^k_{i,h}$ be a binary flag indicating that the data of CH i is successfully received at CH i_{h-1} along the same route to sink k. It follows that, at each particular time slot, we have:

$$Pr\{r^k_{i,h}=1\} = Pr\{r^k_{i,h}=1|t^k_{i,h}=1\}Pr\{t^k_{i,h}=1\}. \tag{3}$$

Consider that a packet of CH i is received at sink k in slot ν. This implies that there exists a scheduling slot vector $\boldsymbol{\nu} = [\nu - \Delta\nu_{N_i^k-1}, ..., \nu - \Delta\nu_1, \nu]$, such that all nodes along the routing path from i to the sink successfully transmit the packet of node i. More specifically, node i_h is scheduled to transmit in slot $\nu - \Delta\nu_{h-1}$, where $\Delta\nu_x > \Delta\nu_y$, $\forall x > y$ and $\Delta\nu_0 = 0$. Along slot vector $\boldsymbol{\nu}$, define the transmission flag of CH i as $t_i^k(\boldsymbol{\nu})$, such that $t_i^k(\boldsymbol{\nu}) = [1, .., 1]$ when CH i transmits a packet to sink k and the transmission at the last hop (at CH i_1) occurs in slot ν. Note that if the relay at the last hop along the transmission path from i to the sink transmits the packet of node i, then it implies that all intermediate hops were scheduled to transmit in prior slots. That is, we have

$$Pr\{t_i^k(\boldsymbol{\nu})=1\} = Pr\{t^k_{i,1}(\nu)=1, ..., t^k_{i,N_i^k}(\nu - \Delta\nu_{N_i^k-1})=1\}. \tag{4}$$

Omit the slot index, (4) can be simplified as: $Pr\{t_i^k=1\} = Pr\{t^k_{i,1}=1, ..., t^k_{i,N_i^k}=1\}$.

For the throughput calculation here, we do not consider retransmissions of packets. Assuming that there exists a schedule such that the source CH and all its intermediate relays are assigned time slots to transmit/forward the source's data, and assuming that the transmissions in all slots are i.i.d, then we can drop the slot index from the throughput expression. In the case when the amplify-and-forward protocol is adopted in the relaying process, which implies that $r^k_{i,h}$'s are independent at different hops, it follows from (1) and (3) that:

$$T_{i,k} = Pr\{t_i^k=1\}\prod_{h=1}^{N_i^k} Pr\{r^k_{i,h}=1|t^k_{i,h}=1\}. \tag{5}$$

The probability of transmission $Pr(t_i^k=1)$, depends on the scheduling scheme adopted in the network. When TDMA protocol is used, each node connected to sink k can transmit with a probability $Pr(t_i^k=1) \geq \frac{1}{N_{intf}\ n_k}$, where n_k is the number of CHs connected to sink k and N_{intf} as the minimum separation between links for bandwidth reuse. That is, when a transmission is made by a CH, other nodes within a distance of $N_{intf}R_c$ from the transmitting CH should remain silent or use another orthogonal channel.

The probability of successful reception, under normal network conditions, can be viewed as a condition on the received signal to interference and noise ratio $SINR$. Assume the SINR threshold for successful reception is γ, then we have $Pr\{r^k_{i,h}=1|t^k_{i,h}=1\} = Pr\{SINR_{i_h,i_{h-1}} > \gamma\}$, $\forall i, h$. In structured networks, the assignment of channels and time slots can be managed to minimize the interference; hence the interference term becomes negligible, and we can write $SINR_{i,j} = \frac{L_{i,j}^{-\beta}P_i}{N_o}$, where $L_{i,j}$ is the distance between nodes i and j, β is the path loss exponent, P_i is the transmitted power, and N_o is the noise power. Assuming that the power is exponentially distributed with mean $\bar{P}$ for all nodes, then by letting $\lambda_{i,h} = \gamma N_o \left[L_{i_h,i_{h-1}}\right]^\beta$ we have

$$\begin{aligned} Pr\{SINR_{i_h,i_{h-1}} > \gamma\} &= Pr\{P_{i_h} > \lambda_{i,h}\} \\ &= \exp\left\{-\gamma\frac{N_o}{\bar{P}}\left[L_{i_h,i_{h-1}}\right]^\beta\right\}. \end{aligned} \tag{6}$$

From the discussion above, we get the following result:

***Theorem** 1: In a multihop MC-WSN network, assuming exponentially distributed transmit powers, the throughput of CH i along a predefined single routing path to sink k is:*

$$T_{i,k} = Pr\{t_i^k=1\}\exp\left\{-\kappa\sum_{h=1}^{N_i^k}\left[L_{i_h,i_{h-1}}\right]^\beta\right\}, \tag{7}$$

where N_i^k is the number of hops in CH i's transmission, $Pr\{t_i^k=1\}$ is the probability that CH i and all its intermediate relaying nodes are scheduled to transmit the data of CH i to sink k, β is the path loss exponent of the channel, $L_{x,y}$ is the distance between nodes x and y, and $\kappa = \gamma\frac{N_o}{\bar{P}}$.

***Remark** 1: It can be seen from* Theorem 1 *that if the hops are equidistant, the throughput will decrease as the number of hops increases. More specifically, when $L_{i_{h-1},i_h} = L$, $\forall h \in \{1, 2, .., N_i^k\}$, we get $T_{i,k} \propto \exp\{-N_i^k\}$. It follows that $\lim_{N_i^k\to\infty} T_{i,k} = 0$. This result justifies our motivation of limiting the number of hops in the topology design and the CCH and RCHs deployment.*

D. Multihop Multipath Routing Case

In the previous subsection, we considered the case when there is a single pre-defined path between a CH and a sink. Note that, in general, the transmission can go through multipaths due to the existence of network diversity. Multipath routing would also increase the security strength against malicious routing attacks as will be illustrated in Section IV. In this section, we formulate the throughput for the multipath case.

***Theorem** 2: Let N be the maximum number of hops from a CH to a sink along any routing path, and $N_i^k(p) \leq N$ is the number of hops from CH i to sink k along path p. Consider that for each hop number $l \in \{1, 2, ..., N\}$, there are $P_{i,l}$ possible l-hop paths from CH i to the sink(s). Let $T(i|N_i^k(p) = l, \mathcal{P}_i^k = p)$ be the throughput that can be achieved along one of the l-hop paths from source i to sink k assuming the path $\mathcal{P}_i^k = p$. Hence, the throughput of node i to sink k can be calculated as:*

$$T_{i,k} = \sum_{l=1}^{N}\sum_{p=1}^{P_{i,l}} T(i|N_i^k(p) = l, \mathcal{P}_i^k = p)\Pr\{\mathcal{P}_i^k = p, N_i^k(p) = l\} \tag{8}$$

Here, l-hop path means a path that consists of l hops. It is noted that $T(i|N_i^k = l, \mathcal{P}_i^k = p)$ can be obtained from *Theorem 1* by substituting $N_i^k = l$, which is the number of hops along the particular path $\mathcal{P}_i^k = p$. The term $\Pr\{\mathcal{P}_i^k = p|N_i^k = l\}$ depends on the routing protocol. It should be emphasized that when multiple routes are enabled, the utilized scheduling protocol, and hence $P(t_i^k = 1)$, could be different than that in the single routing path case.

E. Total Network Throughput

The *network throughput*, Υ, is defined as the average number of packets received successfully from all clusters per unit time.

Let $\mathcal{N}^k$ be the set of CHs that transmit to sink k. Following *Theorems 1 and 2*, the total throughput of the proposed MC-WSN architecture with K RCHs and a CCH can be obtained as:

$$\begin{aligned}\Upsilon &= \sum_{k=0}^{K}\sum_{i\in\mathcal{N}^k} T_{i,k} \\ &= \sum_{k=0}^{K}\sum_{i\in\mathcal{N}^k}\sum_{l=1}^{N}\sum_{p=1}^{\mathcal{P}_{i,l}} T(i|N_i^k=l,\mathcal{P}_i^k=p) \\ &\qquad\qquad \times \Pr\{\mathcal{P}_i^k=p|N_i^k=l\}\Pr\{N_i^k=l\} \\ &= \sum_{k=0}^{K}\sum_{i\in\mathcal{N}^k}\sum_{l=1}^{N}\sum_{p=1}^{\mathcal{P}_{i,l}} p_i^k(p)\exp\left\{-\kappa\sum_{h=1}^{l}\left[L_{i_h^k,i_{h-1}^k}(p)\right]^{\beta}\right\} \\ &\qquad\qquad \times \Pr\{\mathcal{P}_i^k=p|N_i^k=l\}\Pr\{N_i^k=l\}, \qquad (9)\end{aligned}$$

where n_k is the number of nodes connected to sink k, $L_{i_h^k,i_{h-1}^k}(p)$ is the length between CHs i_h^k and i_{h-1}^k along path p, and $p_i^k(p)$ is the transmission probability of CH i along path p to sink k.

IV. Routing Security for MC-WSN

A. Secure Routing Path Selection

In a multihop routing scheme, a compromised node along any routing path can launch routing attack by dropping packets or modifying their contents as they are being forwarded to a sink. Routing attacks could not be completely resolved through cryptographic approaches solely. To combat this type of routing attack, we propose to use multipath diversity in combination with integrity check. More specifically, we propose the following:

- Each node i establishes multiple, preferably disjoint, routes to sink(s). Every packet is transmitted on a randomly selected route. Recall that the probability that node i selects path p to sink k is $\Pr\{\mathcal{P}_i^k = p|N_i^k(p) = l\}$, where $N_i^k(p)$ is the corresponding number of hops along the path. Note that the number of hops in any path should be limited to a pre-specified number to improve the network throughput performance as illustrated in the previous section.
- Each transmitted packet has a sequence number, which is included in the packet header.
- To ensures that packets have not been modified by malicious nodes along the routing paths, data integrity is checked through a cryptographic keyed Hash function. Assume that each source shares a secret key with the sink(s) it communicates with. A Hash function uses the shared key to produces a fixed size "fingerprint" for each packet. The keyed Hash function enables the sink to immediately verify data integrity. Along with the sequence number check, the sink can adjust the routing probabilities ($\Pr\{\mathcal{P}_i^k = p|N_i^k(p) = l\}$) accordingly, then notify the corresponding source. More specifically, we have the following:
(i) If n_p packets received through routing path p are erroneous (checked through the Hash function) then $\Pr\{\mathcal{P}_i^k = p|N_i^k(p) = l\}$ decreases by δ_l. If over multiple observation periods n_p or more packets are received through path p with no errors, then this path is considered reliable and $\Pr\{\mathcal{P}_i^k = p|N_i^k(p) = l\}$ increases by δ_h. δ_l and δ_h are the decrement and increment step sizes, respectively, and they should be tuned along with n_p to take into account errors resulting from possible poor environmental conditions. The tuning of δ_l and δ_h is subject to the constraint that the sum of probability over all possible paths always equals to 1. After sufficient observations, only the reliable routes will be chosen for data delivery.
Note that δ_h could depend on the number of hops along the path. For example, if two paths have shown to be equally reliable, higher δ_h can be used for the path with lower number of hops. That is, given that paths p_1 and p_2 are equally reliable, if $l_1 < l_2$, then choose δ_{h_1} and δ_{h_2} such that: $\Pr\{\mathcal{P}_i^{k_1} = p_1|N_i^{k_1}(p_1) = l_1\} > \Pr\{\mathcal{P}_i^{k_2} = p_2|N_i^{k_2}(p_2) = l_2\}$. That is, we choose the shortest path for throughput maximization. We would analyze the selection of optimal δ_l and δ_h from an information theoretic perspective in future work.
(ii) Similarly, through sequence number check, a packet drop along any of the paths can be detected. The sink consequently decreases the probability that the source transmits over the corresponding path. The sequence number check can be verified at the MA, then notified to the individual sinks (CCH/RCH).

B. Discussions on Throughput

In this subsection, we discuss the effect of routing attacks on the throughput, then show that the proposed secure routing selection approach can lead to an improved performance under malicious attacks.

We start with a single routing path case, where CH i establishes a single path to its (nearest) sink k, i.e., $P_{i,l} = 1$. Note that a routing path is said to be compromised if one or more nodes along this path is compromised. Let an intermediate node at hop level h be compromised to drop or manipulate packets that are being relayed through it. Therefore, the probability of successful reception $Pr\{r_{i,h}^k = 1\} = 0$, and thus, by referring to (5), the overall throughput of node i will be *zero*. On the other hand, since packets routed along a compromised path will not reach the sink, then this means that the average delay per packet is infinite. *In other words, without routing diversity, routing attacks can lead to a zero throughput and infinite delay.*

Let the number of compromised paths between node i and the sink(s) be P_i^c; hence, there are $P_i^b = P_{i,l} - P_i^c$ paths that do not involve any compromised nodes. Following the proposed approach for routing security, if there is at least one path between a CH and a sink has no compromised nodes, i.e., $P_i^b \geq 1$, then a non-zero throughput can be guaranteed for the node.

Let S_i^b be the set of non-compromised paths, and S_i^c be the set of compromised paths. Therefore, we have:

$$\begin{aligned} T_i = \sum_k \sum_{l=1}^{N} \sum_{p \in S_i^b, S_i^c} & T(i|N_i^k(p) = l, \mathcal{P}_i^k = p) \\ & \times \Pr\{\mathcal{P}_i^k = p, N_i^k(p) = l\}. \\ = \sum_k \sum_{l=1}^{N} \sum_{p \in S_i^b} & T(i|N_i^k(p) = l, \mathcal{P}_i^k = p) \Pr\{\mathcal{P}_i^k = p, N_i^k(p) = l\}. \end{aligned} \tag{10}$$

From the equation above, it is clear that to increase the throughput, packets should be transmitted through reliable paths only by choosing $\Pr\{\mathcal{P}_i^k = p | N_i^k(p) = l\} = 0 \; \forall p \in S_i^c$. Note that $\Pr\{\mathcal{P}_i^k = p | N_i^k(p) = l\}$ can be regarded as the weight of choosing path p. Therefore, to maximize the throughput, we need $\sum_{p \in S_i^b} \Pr\{\mathcal{P}_i^k = p | N_i^k(p) = l\} = 1$. Through the keyed hash function and sequence number check, the unreliable paths can be detected and the sink updates $\Pr\{\mathcal{P}_i^k = p, N_i^k(p) = l\}$ by lowering the probabilities that a compromised path is selected. Eventually, we get: $\sum_{p \in S_i^b} \Pr\{\mathcal{P}_i^k = p | N_i^k(p) = l\} = 1$, which improves the throughput as can be seen in (10).

V. Simulation Result

In this section, we illustrate the effect of the number of hops and the malicious attacks on the throughput. Three cases are considered: shortest path routing in the absence of malicious attacks, multipath routing under malicious attacks with and without the secure routing path selection approach. Assume three paths (p_1, p_2, p_3) are available, with number of hops N_m, N_m+1, N_m+2, respectively, and path p_2 is compromised. The throughput versus the number of hops N_m is plotted in Figure 3. Without the secure route selection scheme, all paths are selected with equal probability. When routing path selection is utilized, data is transmitted on the un-compromised paths p_1 and p_2, with higher probability for the shorter path p_1. Here, $\Pr\{p_1, N_m\} = 0.8$ and $\Pr\{p_3, N_m + 2\} = 0.2$. It can be seen that using secure routing path selection, the throughput performance under malicious attack is significantly improved compared to the case when routing path selection approach is not adopted. It is also clear that the throughput decreases as the number of hops increases. This echoes our theoretical analysis.

VI. Conclusions

In this paper, we analyzed the throughput of the MC-WSN architecture under single path and multipath routing. The obtained throughput expressions revealed the impact of the number of hops on the attained throughput performance, and illustrated the trade-off between throughput efficiency and security strength. It was concluded that the shortest path routing improves the throughput. However, under routing attacks, shortest path routing cannot be optimal if compromised. To combat routing attacks, we presented a secure routing path selection approach, and showed that it can improve the network throughput performance.

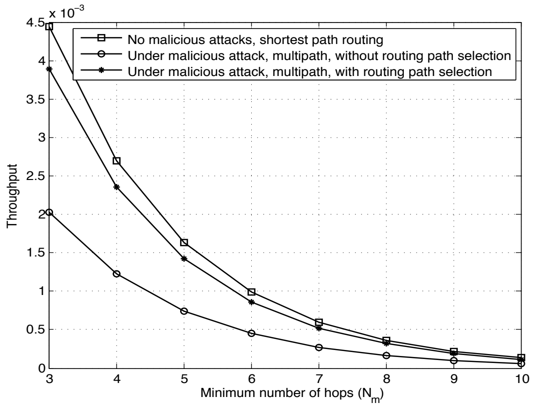

Fig. 3. Throughput versus number of hops. Here, $SINR = \frac{R_c^{-\beta} P_i}{N_o} = 8dB$, $\gamma = 5dB$, $Pr(t_i^k = 1) = 0.02$.

Acknowledgement

This research is partially supported by NSF under awards ECCS-1232109, CNS-1217206 and CNS-0746811.

References

[1] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388 –404, Mar. 2000.

[2] E. J. Duarte-Melo and M. Liu, "Data-gathering wireless sensor networks: organization and capacity," *Computer Networks*, vol. 43, no. 4, pp. 519 – 537, 2003, wireless Sensor Networks. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1389128603003578

[3] M. Grossglauser and D. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Transactions on Networking*, vol. 10, no. 4, pp. 477 – 486, Aug. 2002.

[4] G. Mergen, Z. Qing, and L. Tong, "Sensor networks with mobile access: Energy and capacity considerations," *IEEE Transactions on Communications*, vol. 54, no. 11, pp. 2033 –2044, Nov. 2006.

[5] I. Maza, F. Caballero, J. Capitan, J. Martinez-de Dios, and A. Ollero, "A distributed architecture for a robotic platform with aerial sensor transportation and self-deployment capabilities," *Journal of Field Robotics*, vol. 28, no. 3, pp. 303–328, 2011. [Online]. Available: http://dx.doi.org/10.1002/rob.20383

[6] M. Abdelhakim, J. Ren, J. Ren, and T. Li, "Mobile access coordinated wireless sensor networks – topology design and throughput analysis," *IEEE Global Communications Conference, GLOBECOM'13*, 2013.

[7] C. Karlof and D. Wagner, "Secure routing in wireless sensor networks: attacks and countermeasures," *IEEE International Workshop on Sensor Network Protocols and Applications*, pp. 113 – 127, May 2003.

[8] J. Luo and A. Ephremides, "On the throughput, capacity, and stability regions of random multiple access," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2593 –2607, Jun. 2006.

[9] G. Mergen and L. Tong, "Maximum asymptotic stable throughput of opportunistic slotted ALOHA and applications to CDMA networks," *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1159 –1163, Apr. 2007.

[10] H. Wu, C. Qiao, S. De, and O. Tonguz, "Integrated cellular and ad hoc relaying systems: iCAR," *IEEE Journal on Selected Areas in Communications,*, vol. 19, no. 10, pp. 2105–2115, 2001.

[11] V. Naware, G. Mergen, and L. Tong, "Stability and delay of finite-user slotted ALOHA with multipacket reception," *Cornell University, Ithaca, NY., Technical Report, ACSP-TR-08-04-01*, 2004.