Probabilistic Analysis of Covering and Compromise in Node Capture Attacks

Tamara Bonaci*, Linda Bushnell† and Radha Poovendran*
* Network Security Lab (NSL), EE Dept., University of Washington, Seattle, WA, 98195, USA
† Networked Control Systems Lab, EE Dept., University of Washington, Seattle, WA, 98195, USA

***Abstract*—In this technical report, we analyze wireless sensor networks (WSN) under node capture and cloning attacks. Assuming that WSNs use symmetric keys, with key assignment based on a random key predistribution scheme, we provide extensive probabilistic analysis of WSNs under the attack. We define and characterize the following parameters: the number of nodes an adversary needs to capture in order to disrupt WSN's functionality M, the number of compromised links α due to the fact that an adversary has captured one node, the number of unit refreshment actions after one compromised node is revoked β, the number of valid nodes with all the keys compromised γ and the number of cloned nodes deployed in the network δ.**

I. Preliminaries

In this section, we state the assumptions about the physical WSN system to be analyzed and an adversary performing an attack. A summary of notation used is provided in Table I.

A. Network Model

Consider a WSN containing a set of N sensor nodes, randomly deployed with a density ρ over an area $\mathcal{A}$. We assume WSNs use encrypted communication, with key assignment based on a random key predistribution scheme. Each node is randomly assigned a set of K different keys from a key pool P keys [1]. Two nodes are able to securely communicate if they are within each other's radio range and if they share at least one common key.

Let $\mathcal{N}$ denote a set of N deployed nodes and $\mathcal{K}_t$ a set of symmetric cryptographic keys used for secure communication in a WSN at time t. A WSN can be represented as a random graph $G(\mathcal{N}, \mathcal{K}_t)$, with a set of vertices $\mathcal{N}$ and a set of edges $\mathcal{K}_t$. A pair of nodes $n_i, n_j \in \mathcal{N}$ within each other's radio range is able to securely communicate if and only if they share at least one common key, i.e., $\mathcal{K}_{t,n_i} \cap \mathcal{K}_{t,n_j} \neq \emptyset$.

Let $\mathcal{C}$ denote a set of captured nodes. If there exist a node $c_k \in \mathcal{C}$, a set of keys $\mathcal{K}_{t,c_k}$ held by node c_k is considered to be compromised. Due to the fact that keys are being reused in this predistribution scheme, secure links between any two nodes $n_i, n_j \in \mathcal{N}$ using a key $k_i \in \mathcal{K}_{t,c_k}$ are considered to be exposed to an adversary and hence insecure.

B. Adversarial Model

We consider one active adversary who is assumed to have limited resources and mobility. An adversary is able to actively listen on all of the exposed links throughout the WSN, capture sensor nodes and access all the information stored within them, such as cryptographic keys and measured data. Additionally, an adversary is capable of functionally cloning a captured node and deploying it in a WSN.

TABLE I
A summary of notation used

Symbol	Definition
ρ	Deployment density
$\mathcal{A}$	Area of deployment
$\mathcal{A}_{used}$	Area occupied by valid sensor nodes
$\mathcal{N}$	Set of sensor nodes deployed in the network
R	Radio range of one sensor node
N	Number of nodes in the network
P	Size of key pool
$\mathcal{K}_t$	Set of symmetric cryptographic keys at time t
K	Number of distinct keys assigned to each node
$\mathcal{K}_{t,n_i}$	Set of keys held by the valid node n_i at time t
$\mathcal{K}_{t,n_i \cap n_j}$	Keys nodes n_i and n_j have in common at time t
$\mathcal{C}$	Set of compromised nodes
$\mathcal{K}_{t\mathcal{C}}$	Set of keys held by compromised nodes at time t
$\mathcal{K}_{t,c_k}$	Set of keys held by the compromised node c_k at time t
M	Number of nodes needed to be captured in order to compromise all the links in a WSN
λ_i	Number of nodes sharing the key k_i
α_{max}	Maximum number of compromised links due to existence of a captured node c_k
α_{avg}	Average number of compromised links due to existence of a captured node c_k
β	Number of unit refreshment actions after the revocation of one compromised key $c_k \in \mathcal{C}$
γ	Number of valid nodes with all the keys compromised
δ_i	Maximum number of replicas of one captured node $c_i \in \mathcal{C}$ that can be deployed in a WSN using RM detection algorithm
δ	Maximum number of cloned nodes deployed in a WSN

An adversary's goal is to gain control of a WSN. He achieves that goal by capturing enough nodes to be able to actively listen on all the links used in a WSN or by gathering all the distinct keys assigned to nodes in a WSN.

II. Analysis

In this section we provide the analysis and characterization of parameters M, α, β, γ and δ.

A. Derivation of Parameter M

We start by defining the parameter M, the number of nodes an adversary should capture in order to disrupt network connectivity.

In a WSN using a random key predistribution scheme [1], each node $n_i \in \mathcal{N}$ is assigned a set of K keys. To be able to securely communicate, two nodes n_i and $n_j \in \mathcal{N}$ should be

within each other's radio range and they should have at least one key in common.

Due to the fact that keys are being reused in a WSN, for any key predistribution scheme, there exists an upper bound on number of nodes an adversary needs to capture in order to disrupt network connectivity. For a random key predistribution scheme, the number of nodes a adversary needs to capture in order to disrupt network connectivity, M, is characterized in the following theorem.

Theorem 1: In a WSN using a random key predistribution scheme [1] the number of nodes an adversary needs to capture, M, is equal to

$$M = \frac{P}{K}\left[1 - \left(1 - \frac{K}{P}\right)^N\right] \tag{1}$$

Proof: As noted in [2], a single key $k_j \in \mathcal{K}_t$ is assigned to the node $n_i \in \mathcal{N}$ with the probability $\frac{K}{P}$, which can be modeled as Bernoulli random variable.

Under the assumption that each node $n_i \in \mathcal{N}$ is independently assigned a set of keys, the process of assigning the key $k_j \in \mathcal{K}_t$ can be modeled as binary switch process and probability that the key $k_j \in \mathcal{K}_t$ is not assigned to any node can be calculated as:

$$\mathbb{P}[\text{key } k_i \text{ is not assigned to any node}] = \left(1 - \frac{K}{P}\right)^N = \hat{p}$$

Now the expected number of keys assigned to at least one node can be calculated as:

$$\begin{aligned}\mathbb{E}[K_{asg}] &= P\mathbb{E}[\text{key } k_i \text{ assigned to at least one node}] \\ &= P\left[1 - \hat{p}\right] = P\left[1 - \left(1 - \frac{K}{P}\right)^N\right]\end{aligned}$$

Assuming that there exists a subset of nodes $\mathcal{N}_{max} \in \mathcal{N}$ in the WSN such that any two nodes n_i and $n_j \in \mathcal{N}_{max}$ do not share a key (maximum non-overlapping set of nodes), the minimum number of nodes an adversary needs to capture is defined as:

$$M = \frac{P}{K}\left[1 - \left(1 - \frac{K}{P}\right)^N\right]$$

■

Equation (1) represents the lower bound on the number of nodes an adversary needs to capture in order to disrupt network connectivity, since the analysis is based on the existence the maximum non-overlapping set of nodes $\mathcal{N}_{max}$. It is however known that capturing the maximum non-overlapping set of nodes is NP-hard [3].

B. Derivation of Number of Compromised Links

Consider a situation of one compromised node $c_k \in \mathcal{C}$ in a WSN that holds exactly K distinct keys. Since an adversary is assumed to actively listen on all of the exposed links, i.e., all the links he holds the keys for, the communication on all such links is considered to be broken. Therefore, the impact of one compromised node on the connectivity of WSN can be analyzed in terms of the number of compromised links. We define the parameter α as the number of compromised links and analyze two situations: the maximum number of broken links and the average number of broken links.

In order to characterize parameters α_{max} and α_{avg}, let's first define the parameter λ_i, representing the number of nodes sharing a key $k_i \in \mathcal{K}_t$

Lemma 1: The probability distribution of the parameter λ_i, representing the number of nodes sharing a key $k_i \in \mathcal{K}_t$ can be modeled as the binomial distribution $\mathcal{B}(N, \frac{K}{P})$ [2]:

$$\mathcal{P}(\lambda) = \binom{N}{\lambda}\left(\frac{K}{P}\right)^{\lambda}\left(1 - \frac{K}{P}\right)^{N-\lambda} \tag{2}$$

Proof: In a WSN with N nodes, using a random key predistribution scheme [1], each node $n_i \in \mathcal{N}$ is randomly assigned a set of K distinct keys. A particular key $k_j \in \mathcal{K}_t$ is selected with probability $\frac{K}{P}$, which can be modeled as Bernoulli random variable [2].

Under the assumption that each node $n_i \in \mathcal{N}$ is independently assigned a set of keys $\mathcal{K}_{t,n_i}$ from the key pool of P distinct keys, the probability distribution $\mathcal{P}(\lambda)$ that the key $k_j \in \mathcal{K}_t$ is shared by exactly λ_j nodes can be modeled as a binomial distribution:

$$\mathcal{P}(\lambda) = \binom{N}{\lambda}\left(\frac{K}{P}\right)^{\lambda}\left(1 - \frac{K}{P}\right)^{N-\lambda}$$

with expected value and variance:

$$\begin{aligned}\mathbb{E}[\lambda] &= \frac{NK}{P} \\ \sigma_\lambda^2 &= \frac{NK^2}{P^2}\left(\frac{P}{K} - 1\right)\end{aligned}$$

■

Theorem 2: The maximum number of exposed links due to a compromised node c_k is equal to

$$\alpha_{max} = \binom{N}{2}\left(\frac{K^3}{P^2}\right) \tag{3}$$

Proof: Consider a single key $k_j \in \mathcal{K}_t$. There exist λ_i nodes in a WSN sharing a key k_i. Hence, there can exist at most $\binom{\lambda_i}{2}$ communication links formed using the key k_i.

Assuming that the maximum number of links is formed using every key $k_i \in \mathcal{K}_t$, the maximum number of links exposed to an adversary by capturing the node $c_k \in \mathcal{C}$ is equal to the sum

$$\sum_{i \in \mathcal{K}_{t,c_k}} \binom{\lambda_i}{2}$$

The maximum number of links exposed to an adversary can be calculated as:

$$\begin{aligned}\alpha_{max} &= \mathbb{E}\left[\sum_{i \in \mathcal{K}_{t,c_k}} \binom{\lambda_i}{2}\right] \\ &= \sum_{i \in \mathcal{K}_{t,c_k}} \mathbb{E}\left[\frac{\lambda_i^2 - \lambda_i}{2}\right]\end{aligned} \tag{4}$$

Equation (4) can be rewritten, by noting that each node $n_j \in \mathcal{N}$ holds exactly K keys:

$$\alpha_{max} = \frac{K}{2}(\mathbb{E}[\lambda_i^2] - \mathbb{E}[\lambda_i]) \tag{5}$$

By noting that λ_i is binomial random variable (equation (2)), equation (5) can be rewritten as:

$$\alpha_{max} = K\left(\frac{K^2}{P^2}\right)\frac{N(N-1)}{2} = \binom{N}{2}\frac{K^3}{P^2} \tag{6}$$

■

Theorem 3: The average number of exposed links due to a compromised node c_k is equal to:

$$\alpha_{avg} = \binom{N}{2} p\left\{1 - \left[1 - \frac{K}{P}\frac{N\left(\frac{K}{P}\right)-1}{N-1}\right]^K\right\}$$

where p represents the probability that any two nodes in a WSN share a link.

Proof: In order to derive the parameter α_{avg}, let's recall the random graph representation of a WSN $G(\mathcal{N}, \mathcal{K}_t)$, where the set of sensor nodes $\mathcal{N}$ represents the set of graph vertices and the key pool $\mathcal{K}_t$ the set of edges at time t.

Now the average number of links node $c_k \in \mathcal{C}$ shares with other nodes in a WSN can be expressed as:

$$Z = \sum_{(j,l)\in(\mathcal{N}\times\mathcal{N})} \mathbb{1}\{(j,l) \in \mathcal{K}_t \text{ and } \mathcal{K}_{t,j\cap l} \cap \mathcal{K}_{t,c_k} \neq \emptyset\}$$

where $\mathbb{I} = \mathbb{1}\{(j,l) \in \mathcal{K}_t \text{ and } \mathcal{K}_{t,j\cap l} \cap \mathcal{K}_{t,c_k} \neq \emptyset\}$ represents the indicator function:

$$\mathbb{I} = \begin{cases} 1, & \text{if } \left(\mathcal{K}_{t,n_j} \cap \mathcal{K}_{t,n_l}\right) \cap \mathcal{K}_{t,c_k} \neq \emptyset \\ & \quad \text{and } \exists \text{ a link between the nodes } n_j \text{ and } n_l \\ 0, & \text{otherwise} \end{cases}$$

The expected value of the average number of links node c_k shares with other nodes can now be calculated as:

$$\mathbb{E}[Z] = \sum_{(j,l)\in(\mathcal{N}\times\mathcal{N})} \mathbb{P}[(j,l) \in \mathcal{K}_t \text{ and } \mathcal{K}_{t,j\cap l} \cap \mathcal{K}_{t,c_k} \neq \emptyset] \tag{7}$$

Under assumption that node deployment and key assignment are independent processes, equation (7) can be rewritten as:

$$\mathbb{E}[Z] = \sum_{(j,l)\in(\mathcal{N}\times\mathcal{N})} \mathbb{P}[\mathcal{K}_{t,j\cap l} \cap \mathcal{K}_{t,c_k} \neq \emptyset]\mathbb{P}[(j,l) \in \mathcal{K}_t] \tag{8}$$

where $p = \mathbb{P}[(j,l) \in \mathcal{K}_t]$ denotes the probability that two nodes $n_j, n_l \in \mathcal{N}$ share a link and $\mathbb{P}[\mathcal{K}_{t,j\cap l} \cap \mathcal{K}_{t,c_k} \neq \emptyset]$ denotes the probability that two nodes $n_j, n_l \in \mathcal{N}$ share at least one common key with the captured node $c_k \in \mathbb{C}$.

By assumption that existences of each link (j,l) in the set of edges $\mathcal{K}_t$ are independent, identically distributed (i.i.d) random variables, equation (8) can be rewritten as:

$$\mathbb{E}[Z] = \binom{N}{2} p(1 - \mathbb{P}[\mathcal{K}_{t,j\cap l} \cap \mathcal{K}_{t,c_k} = \emptyset]) \tag{9}$$

The probability $\mathbb{P}[\mathcal{K}_{t,j\cap l} \cap \mathcal{K}_{t,c_k} = \emptyset]$ can be calculated as:

$$\begin{aligned} & \mathbb{P}[\mathcal{K}_{t,j\cap l} \cap \mathcal{K}_{t,c_k} = \emptyset] \\ = & \sum_{i=1}^{K} \mathbb{P}[|\mathcal{K}_{t,j\cap l}| = i, \mathcal{K}_{t,j\cap l} \cap \mathcal{K}_{t,c_k} = \emptyset] \\ = & \sum_{i=1}^{K} \mathbb{P}[|\mathcal{K}_{t,j\cap l}| = i]\mathbb{P}[\mathcal{K}_{t,j\cap l} \cap \mathcal{K}_{t,c_k} = \emptyset] \end{aligned} \tag{10}$$

with probability $\mathbb{P}[|\mathcal{K}_{t,j\cap l}| = i]$ defined as [1]:

$$\begin{aligned} \mathbb{P}_1 &= \mathbb{P}[|\mathcal{K}_{t,j\cap l}| = i] \\ &= \mathbb{P}[\text{nodes } n_j, n_l \text{ share exactly } i \text{ links}] \\ &= \binom{K}{i}\left(\frac{\lambda-1}{N-1}\right)^i\left(\frac{N-\lambda}{N-1}\right)^{K-i} \end{aligned}$$

and the probability $\mathbb{P}[\mathcal{K}_{t,j\cap l} \cap \mathcal{K}_{t,c_k} = \emptyset]$ as:

$$\begin{aligned} \mathbb{P}_2 &= \mathbb{P}[\mathcal{K}_{t,j\cap l} \cap \mathcal{K}_{t,c_k} = \emptyset] \\ &= \mathbb{P}[\nexists \text{ common key } k_i \text{ between the nodes } n_j, n_l, c_k] \\ &= \left(1 - \frac{K}{P}\right)^{|\mathcal{K}_{t,n_j\cap n_l}|} \end{aligned} \tag{11}$$

Equation (11) is derived based on the assumption that each key $k_i \in \mathcal{K}_t$ is independently assigned to a node $n_j \in \mathcal{N}$ with the probability $\frac{K}{P}$.

Now equation (10) can be rewritten as:

$$\begin{aligned} & \mathbb{P}[\mathcal{K}_{t,j\cap l} \cap \mathcal{K}_{t,c_k} = \emptyset] \\ = & \sum_{i=1}^{K}\left(1 - \frac{K}{P}\right)^i\binom{K}{i}\left(\frac{\lambda-1}{N-1}\right)^i\left(\frac{N-\lambda}{N-1}\right)^{K-i} \\ = & 1 - \left(\frac{K}{P}\right)\frac{N\left(\frac{K}{P}\right)-1}{N-1} \end{aligned}$$

Finally, equation (9) can be written as:

$$\mathbb{E}[Z] = \binom{N}{2} p\left\{1 - \left[1 - \left(\frac{K}{P}\right)\frac{N\left(\frac{K}{P}\right)-1}{N-1}\right]^K\right\}$$

■

Remark 1: The probability that any two nodes $n_i, n_j \in \mathcal{N}$ in the WSN share a link, p, can be calculated as in [1]:

$$p = 1 - \frac{\left(1 - \frac{K}{P}\right)^{2\left(P-K+\frac{1}{2}\right)}}{\left(1 - \frac{2K}{P}\right)^{\left(P-2K+\frac{1}{2}\right)}} \tag{12}$$

C. Derivation of Number of Unit Refreshment Actions

When a cloned node $c_k \in \mathcal{C}$ is detected in a WSN, such a node, as well as all the cryptographic keys held by that node, are revoked. Due to the fact that cryptographic keys are being reused in a WSN, there exists a non-zero probability that a pair of valid nodes $n_i, n_j \in \mathcal{N}$ used one of the revoked keys $k_l \in \mathcal{K}_{t,c_k}$. In order to maintain network connectivity, the revocation action in a WSN is followed by the key refreshment action, during which all the nodes holding at least one revoked key $k_l \in \mathcal{K}_{t,c_k}$ are being updated with freshly generated keys. We define the parameter β as the number of unit refreshment actions due to the fact that cloned node $c_k \in \mathcal{C}$ was revoked. The unit refreshment action represents the update of one node $n_i \in \mathcal{N}$ with one fresh key $k_m \in \mathcal{K}_t$.

Theorem 4: In a WSN, using a random key predistribution scheme, the number of unit refreshment actions, due to the fact that cloned node $c_k \in \mathcal{C}$ is revoked, β, is equal to:

$$\beta = \frac{K^2 N}{P} \tag{13}$$

Proof: Consider a WSN with one compromised node $c_k \in \mathcal{C}$. There exist λ_i distinct nodes sharing each of the compromised keys $k_l \in \mathcal{K}_{t,c_k}$. As shown in equation (2), the parameter λ_i is binomial random variable with average value, defining the expected number of nodes holding each key, equal to:

$$\mathbb{E}[\lambda] = N\frac{K}{P}$$

Since each of K revoked keys k_i should be sent to $N\frac{K}{P}$, the number of unit refreshment actions is equal to:

$$\beta = \frac{K^2 N}{P}$$

■

D. Derivation of Number of valid nodes with all keys compromised

Due to the fact that keys are being reused in a WSN, there exists a non-zero probability that the union of the sets of keys held by all the compromised nodes $\mathcal{K}_{t,\mathcal{C}}$ completely covers the set of keys assigned to one or more valid nodes $n_i \in \mathcal{N}$. Valid nodes with all the keys compromised are not able to securely communicate. We define the parameter γ as the number of valid nodes with all the keys compromised and characterize it as follows.

Theorem 5: The number of valid nodes with all the keys compromised γ is a binomial random variable $\mathcal{B}([N-|\mathcal{C}|], p^*)$, where p^* denotes the probability that all of the keys of the valid node $n_i \in \mathcal{C}$ are compromised and is defined as:

$$p^* = \left(1 - \left[\frac{N-\lambda_i}{N-1}\right]^{|\mathcal{C}|}\right)^K \tag{14}$$

where $|\mathcal{C}|$ denotes the number of compromised keys.

Proof: Consider a node $n_i \in \mathcal{N}$, holding a set of keys $\mathcal{K}_{t,n_i}$, and a set of compromised nodes $\mathcal{C}$, holding a set of compromised keys $\mathcal{K}_{t,\mathcal{C}} = \cup_{k=1}^{|\mathcal{C}|}\mathcal{K}_{t,c_k}$.

The probability that the node n_i has at least one valid key can be calculated as:

$$\mathbb{P}[n_i\text{holds at least one valid key}] = 1 - \mathbb{P}[n_i\text{has no valid nodes}]$$

Under assumption that each key $k_l \in \mathcal{K}_{t,n_i}$ is assigned independently to the node n_i, the probability that the node n_i does not have valid keys can be calculated as follows:

$$\mathbb{P}[n_i \text{ has no valid keys}] = p^* = \prod_{l=1}^{K}\mathbb{P}[k_l \subset \mathcal{K}_{t,\mathcal{C}}]$$

By assumption that one node $c_k \in \mathcal{C}$ is captured independently of other nodes, probability p^* can be rewritten as:

$$p^* = \prod_{l=1}^{K}\left[1 - \prod_{k=1}^{|\mathcal{C}|}\mathbb{P}\left[k_l \not\subset \mathcal{K}_{t,c_k}\right]\right]$$

Since the key k_l is held by λ_l nodes, the probability $\mathbb{P}[k_l \not\subset \mathcal{K}_{t,c_k}]$ is equal to $\frac{N-\lambda_l}{N-1}$. Therefore, the probability p^* can be written as:

$$\begin{aligned} p^* &= \prod_{l=1}^{K}\left[1 - \prod_{k=1}^{|\mathcal{C}|}\mathbb{P}\left[k_l \not\subset \mathcal{K}_{t,c_k}\right]\right] \\ &= \prod_{l=1}^{K}\left[1 - \left(\frac{N-\lambda_l}{N-1}\right)^{|\mathcal{C}|}\right] \end{aligned}$$

Since all of the parameters λ_l are identically distributed, the probability p^* can be written as:

$$p^* = \left[1 - \left(\frac{N-\lambda_l}{N-1}\right)^{|\mathcal{C}|}\right]^K$$

Finally, the parameter γ can be defined as binomial random variable, with mean and variance:

$$\begin{aligned} \mathbb{E}[\gamma] &= Np^* = N\left[1 - \left(\frac{N-\lambda_l}{N-1}\right)^{|\mathcal{C}|}\right]^K \\ &= N\left[1 - \left(\frac{N\left(1-\frac{K}{P}\right)}{N-1}\right)^{|\mathcal{C}|}\right]^K \\ \sigma_\gamma^2 &= Np^*(1-p^*) \end{aligned}$$

■

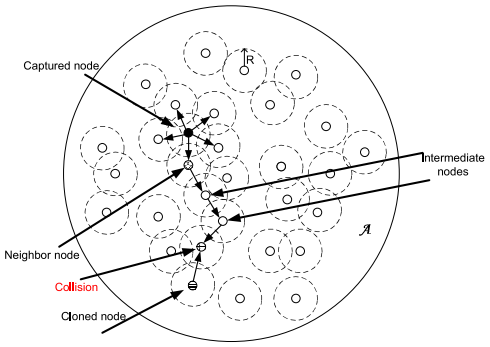

Fig. 1. Line-Select Multicast cloned nodes detection algorithm: In a WSN with N nodes, deployed over an area $\mathcal{A}$ there exists one captured node and one replica of that node. Each node has the same radio range R and periodically broadcasts its location and ID to its neighbor nodes. Neighbor nodes randomly chose a set of witness nodes to send a broadcasted message to. If any of the intermediate nodes detects a node with the same ID, but different location, a collision occurs and the cloned node is detected.

E. Derivation of Number of Cloned Nodes

In addition to being able to actively listen on all of the compromised links, to capture a node and to steal all the information within that node, an adversary is able to functionally clone a captured node and deploy it in a WSN. An adversary, however, performs the process of cloning and deploying cloned nodes in a WSN with caution, in order to avoid detection of cloned nodes, since there exist efficient detection algorithms. We define the parameter δ as the number of cloned nodes deployed in a WSN, when a WSN uses cloned nodes detection algorithms, such as Random Multicast (RM) or Line-Select Multicast(LSM) [4], in order to detect cloned nodes (Fig 1).

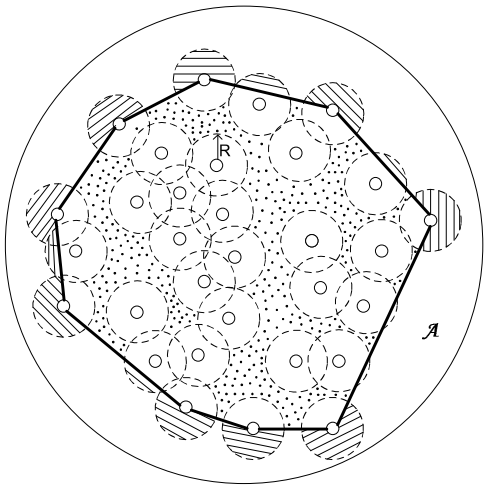

Fig. 2. The expected area occupied by N nodes. Area is calculated as a convex hull defined by N nodes. It is assumed that the crossed area, representing the radio range areas excluded from the convex hull and dotted area, representing empty area, i.e., an area not occupied by nodes or within radio range of any deployed node, are approximately equal.

Prior to characterizing the parameter δ, let's define the parameter $\mathcal{A}_{used}$ as the expected area occupied by deployed valid nodes.

Theorem 6: The expected area (Figure 2) occupied by N deployed valid nodes is equal to:

$$\mathcal{A}_{used} = \mathcal{A}\frac{N!}{4\pi}\sum_{k=0}^{[(N-3)/2]}\frac{(-1)^k}{(2\pi)^{2k}(N-2k-2)!} \tag{15}$$

Proof: Consider a WSN with N sensor nodes deployed over an area $\mathcal{A}$. Assuming that sensor nodes are deployed uniformly and independently over a unit disc, the convex hull defined by the sensor nodes has the area equal to [5]:

$$A^* = \frac{N!}{4\pi}\sum_{k=0}^{[(N-3)/2]}\frac{(-1)^k}{(2\pi)^{2k}(N-2k-2)!}$$

Now let's consider a case when sensors are not deployed over a unit ball, but over a ball of area $\mathcal{A}$. The area occupied by N sensors can then be defined as:

$$\mathcal{A}_{used} = \mathcal{A}\frac{N!}{4\pi}\sum_{k=0}^{[(N-3)/2]}\frac{(-1)^k}{(2\pi)^{2k}(N-2k-2)!}$$

■

In order to define the parameter δ, we define the parameter δ_i, representing the maximum number of replicas of one captured node $c_i \in \mathcal{C}$ that can be deployed in a WSN using cloned nodes detection algorithms, such as RM or LSM [4].

Theorem 7: The maximum number of replicas $\tilde{c}_i$ of a captured node $c_i \in \mathcal{C}$ in a WSN using clone detection algorithm, such as RM, is bounded above by the following expression:

$$\delta_i = \frac{N!}{16R^2\pi^2}\sum_{k=0}^{[(N-3)/2]}\frac{(-1)^k}{(2\pi)^{2k}(N-2k-2)!} - 1 \tag{16}$$

Proof: Let's consider the situation where an adversary captures one sensor node $c_i \in \mathcal{C}$ and tries to deploy as much as possible exact replicas $\tilde{c}_i$ of the captured node in a WSN.

A naive adversary could try to deploy clones $\tilde{c}_i$ randomly inside an area of deployment $\mathcal{A}$. Such an approach however wouldn't be very efficient, since we are assuming WSNs are using cloned nodes detection algorithms.

We therefore conclude that for each captured node c_i there exists an area $\mathcal{A}_{off}$, with the probability that an adversary will deploy cloned nodes within that area is equal to zero.

The area with zero probability of cloned node deployment is characterized by the radio range of the captured node $c_i \in \mathcal{C}$, R, and its intermediate neighbors and can be approximated as:

$$\mathcal{A}_{off} = 4R^2\pi \tag{17}$$

Now we define an area were an adversary can randomly deploy clones of the captured node $c_i \in \mathcal{C}$ as:

$$\begin{aligned}\mathcal{A}_{allowed} &= \mathcal{A}_{used} - A_{off} \\ &= \mathcal{A}\frac{N!}{4\pi}\sum_{k=0}^{[(N-3)/2]}\frac{(-1)^k}{(2\pi)^{2k}(N-2k-2)!} - 4R^2\pi\end{aligned}$$

Due to the fact that cloned nodes detection algorithms are used, we know that an adversary will never deploy new replica of captured node c_i inside the radio range of existing replica or one-hop neighbors of that replica. Therefore the upper bound on the number of replicas of the same captured node c_i an adversary can deploy can be characterized as:

$$\begin{aligned}\delta_i &= \frac{\mathcal{A}_{allowed}}{4R^2\pi} \\ &= \frac{N!}{16R^2\pi^2}\sum_{k=0}^{[(N-3)/2]}\frac{(-1)^k}{(2\pi)^{2k}(N-2k-2)!} - 1\end{aligned}$$

■

Theorem 8: The maximum number of cloned nodes δ an adversary is able to deploy in a WSN using cloned nodes detection algorithm, such as RM, is equal to:

$$\begin{aligned}\delta &= M \cdot \delta_i \\ &= M\frac{N!}{4R^2\pi^2}\sum_{k=0}^{[(N-3)/2]}\frac{(-1)^k}{(2\pi)^{2k}(N-2k-2)!} - M\end{aligned} \tag{18}$$

where M denotes the maximum number of nodes an adversary needs to capture in order to disrupt the connectivity of a WSN and is defined with equation (1).

Proof: Consider a WSN with N nodes deployed over an area $\mathcal{A}$, using clone detection algorithm, such as RM.

The maximum number of replicas of one captured node $c_i \in \mathcal{C}$ that can be deployed in a WSN is characterized by equation (1). If we assume that deployments of replicas of distinct captured nodes $c_i, c_j \in \mathcal{C}$ are independent, the maximum number of the cloned nodes an adversary can deploy is equal to:

$$\begin{aligned}\delta &= M \cdot \delta_i \\ &= M\frac{N!}{4R^2\pi^2}\sum_{k=0}^{[(N-3)/2]}\frac{(-1)^k}{(2\pi)^{2k}(N-2k-2)!} - M\end{aligned}$$

where M comes based on the assumption that an adversary will never try to capture more than M distinct nodes. ■

Remark 1: From Theorem 7 it follows that, once replica is deployed in a WSN, there is no difference between the replica $\tilde{c}_i$ and the original captured node $c_i \in \mathcal{C}$ in terms of deploying a new replica of the same node c_i.

REFERENCES

[1] L. Eschenauer and V. D. Gligor, "A key-management scheme for distributed sensor networks," in *In Proceedings of the 9th ACM Conference on Computer and Communications Security*. ACM Press, 2002, pp. 41–47.

[2] P. Tague and R. Poovendran, "A canonical seed assignment model for key predistribution in wireless sensor networks," *TOSN*, vol. 3, no. 4, 2007.

[3] ——, "Modeling adaptive node capture attacks in multi-hop wireless networks," *Ad Hoc Networks*, vol. 5, no. 6, pp. 801 – 814, 2007.

[4] B. Parno, A. Perrig, and V. D. Gligor, "Distributed detection of node replication attacks in sensor networks," in *IEEE Symposium on Security and Privacy*, 2005, pp. 49–63.

[5] A. M. Mathai, *Introduction to Geometrical Probabillity: Distributional Aspects with Applications*. The Netherlands: Gordon and Breach, 1999.