CKEF: A Cluster-based Key Establishment Framework for homogenous mobile and static wireless sensor networks

Mohammad Rezaeirad[§], Sahar Mazloom[§], Mahdi Orooji[†], Miao Jin[§], and Magdy Bayoumi[§]
[§]The Center for Advanced Computer Studies, University of Louisiana at Lafayette, Lafayette, LA, USA
Email: {rxm1725, sxs7444, mjin, mab}@cacs.louisiana.edu
†Department of Electric and Computer Engineering, Louisiana State University, Baton Rouge, LA, USA
Email: {morooj1}@lsu.edu

***Abstract*—Mission critical applications on homogenous mobile wireless sensor networks (HMWSNs) mandate new sets of security appliances to be friendly with existing resource constrained hardware platforms. To deliver a promising security, particularly in military deployments, mechanisms have to build upon an efficient key management that compensates HMWSNs constraints. Cluster-based key establishment is being the prime focus among the recent works in key establishment due to its significant improvement on network efficiency, security, scalability and flexibility. Therefore, we propose a Cluster-based framework to support pre-distribution key establishment schemes for HMWSNs. The proposed framework is compatible with most of pre-distribution schemes, and two instantiations are provided in this work to support our claim that the proposed framework improves security and scalability of the adopted schemes. We develop analytical models and conduct extensive simulations to evaluate the security and performance of the proposed framework, and the network connectivity under different scenarios.**

I. Introduction

Homogeneous mobile wireless sensor networks (HMWSNs) deploy hundreds, even thousands of sensor nodes with identical hardware specifications in a range of domains for varieties of mission critical applications. These applications are resilient to attacks, but the hostility of their deployment environments does not allow maintenance after node installation [1]. Therefore, security countermeasures should be thoroughly designed prior to these missions. Cryptographic appliances are promising tools to safeguard the sensor nodes against most of known attacks, and key management is a basis for most cryptosystems.

Considering the power and storage constraints of each sensor node in HMWSNs, a Key establishment should have a lightweight design when satisfying the security requirements. The well accepted energy-efficient approach for key establishment in HMWSNs is based on the idea of the distribution of key before the deployment of a network [2], [3]. Considering security issue and scalability limitation imposed by key pre-distribution [4], an alternative approach is to pre-load sensors with a key generation method introducing limited computational overhead instead of the actual keys [5], [6]. Blundo's scheme [6] relies on pre-distribution of polynomial shares of a randomly generated and symmetric bivariate $t-$degree polynomial. The major limitation of Blundo's scheme is the lack of scalability because the security of a network will be jeopardized after the capture of $t+1$ number of nodes [7]. Works proposed in [8]–[10] utilize a set of perturbation polynomials to solve the scalability problem of Blundo's scheme by increasing the resiliency threshold while maintaining efficiency. Unfortunately, a comprehensive security argument given in [11] indicates that these modified schemes can be broken easily by attackers. Later, Liu and Ning [12] apply deployment knowledge to enhance the scalability and connectivity of q-composite scheme [3], and combine with Blundo's scheme to achieve higher connectivity and less communication overhead.

Unfortunately, scalability is still an unsolved issue of the discussed pre-distribution schemes. A wireless sensor network can be divided to clusters. Cluster-based pre-distribution schemes improve network security, scalability, and flexibility. Specifically, cluster keying offers more resilience based on the fact that compromising a node impacts only residing cluster rather than the entire of a network. Node addition and revocation become more flexible for a large network and this makes the design of a scalable pre-distribution scheme possible. However, most of previous cluster-based pre-distribution schemes consider heterogeneous wireless sensor networks only and assume cluster-heads with stronger hardware capacity (better computation power, memory storage, and radio coverage). Cluster-heads have better control against security attacks, and sometimes they are even assumed to be totally secured. These assumptions are not practical for many mission critical applications such as military espionage operations to detect moving targets or the attendance of micro agent listeners [13]. However, only a few previous works consider a cluster-based key establishment scheme on homogenous wireless sensor networks in the last few years [14]–[19]. They assume deployment knowledge available and consider only static sensor nodes.

Therefore, we propose a cluster-based framework to support pre-distribution key establishment schemes for HMWSNs. In

the proposed framework, the number of clusters is determined by the specific requirements of target applications of a deployed network. Sensor nodes are grouped to clusters based on their pre-loaded key materials. Nodes with the same key materials can establish key for intra-cluster communication. Based on application necessities, a small number of nodes are randomly chosen as cluster-heads before network deployment. They are pre-loaded with extra key materials which allow them to establish keys with other cluster-heads for inter-cluster communication to exchange control messages. All the assignments are totally off-line, and a cluster-head does not carry an extra duty in key establishment or discovery for any other nodes. So we don't require a cluster-head armed with stronger hardware specifications to preserve a normal life time. We consider mobility of sensor nodes in the framework and assume nodes may move from a location to another for an assignment. By incorporating a mobile model to this framework, we show that a desired global connectivity can be guaranteed if a required local connectivity is satisfied.

The major contribution of this paper is a proposed cluster-based framework to support key establishment schemes for HMWSNs. The framework can embrace general pre-distribution schemes as its key establishment protocol and support mobile wireless sensor networks (WSNs).

The rest of the paper is organized as follows. Section II gives a brief review of related works. Section III presents the proposed framework in details. Section IV gives two instantiations of the framework. Section V analyzes the performance of the framework. Numerical results are presented in Section VI. Section VII concludes the paper.

II. Related Works

Key management on WSNs has been discussed extensively in previous works [20]–[24]. We only give a brief review of cluster-based key management which is most related to ours. For a complete list of recent cluster-based key management schemes, we refer readers to [13].

Du *et al.* [14] propose a group-based key management scheme by applying deployment knowledge to the basic scheme [2]. Resulting scheme, is more memory efficient and offers less communication overhead as well as a better resilience. Their scheme uses deployment knowledge as an enhancement mechanism, which is not flexible for verities of applications. Similarly, the hexagonal group-based key management [15] combines Blundo's scheme [6] with deployment knowledge. Liu *et al.* [17] introduce a group-based deployment scheme for homogeneous wireless sensor networks. Their framework is compatible with most of the mentioned schemes and help improve the security, performance, and scalability of these underlying schemes. Martin *et al.* [18] improve the flexibility of [17] without increasing storage requirements or sacrificing resilience. All these mentioned cluster-based (group-based) schemes consider using deployment knowledge in their assumption to improve security.

Heterogeneous cluster-based key managements usually assume cluster-heads with stronger hardware specification than other nodes, because cluster-heads may need to carry special duties in key establishments. Lu *et al.* [25] propose a unified framework for distributed key management schemes in heterogeneous wireless sensor networks. The security and performance of the framework highly rely on such stronger cluster heads. The main disadvantage with heterogeneous schemes is that they are not flexible with general network topologies and applications [13], [21], [24]. The assumption taken in many heterogeneous schemes [26]–[28] that cluster-heads are well connected and always secure against attacker, is not practical.

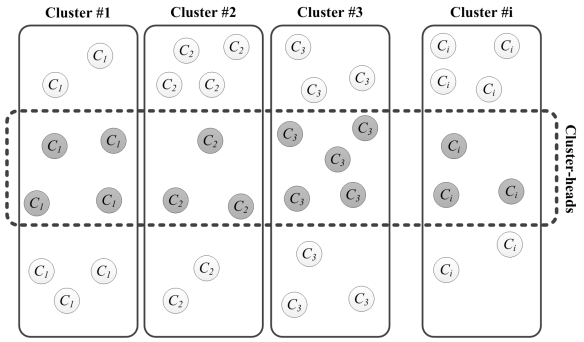


Fig. 1. Conceptual schematic of proposed framework

III. Proposed Framework

In this section, we introduce the important factors of the proposed Cluster-based Key Establishment Framework (CKEF): network elements, attacker properties, cluster model, and key establishment essentials.

A. Network elements

The proposed framework employs a large number of mobile sensor nodes with identical physical characteristic (homogenous). We assume nodes are uniformly deployed in the operation zone. Also, a sensor node has enough neighbors, thus, there would be no isolated nodes. In this decentralized network, a base station could be mobile and operates in a location where it is connected to sensor nodes.

B. Attacker properties

We assume an attacker has a very powerful hardware platform. An attacker can eavesdrop any encrypted or unencrypted conversation between nodes instantly (but needs keys to interpret them), and physically capture nodes and immediately discover their containing information. An attacker tries to capture more nodes from network intangibly, although the attacker has limited time to utilize captured information before victimized node is revoked.

C. Cluster model

Clusters are formed abstractly; nodes group into clusters based on the roles in their assignment of tasks instead of deployment information and geographic information of nodes. Key materials are uniquely designed for each cluster; therefore

clusters provide boundaries for information flow such that there would be no information leakage from one cluster to another. A node cannot leave its current cluster and join another one freely, but the framework provides mechanisms for both node addition and revocation based on adopted underlying schemes. Heads of a cluster are randomly chosen from nodes belonging to the cluster; this confirms that, we do not suggest a heterogeneous design. Fig. 1 illustrates the cluster formation. Communication of cluster-heads (intra-cluster communication) is the only way for two clusters to exchange critical information (such as control messages) in a secure fashion. In some scenario cluster-heads may exchange news of an attack. We assume the amount of inter-cluster communication is much less than intra-cluster. Additionally, two nodes from two different clusters may not have privilege (required keys) to establish a secure connection via heads.

D. Key establishment essentials

The key establishment essentials comprise of three main procedures: pre-distribution, pair discovery and pair-wise key establishment, and path-key establishment. In fact, any key establishment supporting these essentials can fit into the proposed framework. These procedures make it possible to have two levels of communication of Intra-cluster connectivity and Inter-cluster connectivity.

1) Pre-distribution: A large pool of key materials is generated in this off-line process. Each sensor node is loaded with a pre-defined amount of key materials according to the underlying adopted scheme prior to deployment. This procedure make it possible to have two forms of communication of Intra-cluster connectivity and Inter-cluster connectivity.

Intra-cluster pre-distribution: Nodes belonging to a cluster receive a pre-defined amount of key materials from the same pool allocated to that cluster. This allocation enforces a paradigm where nodes from a particular cluster may be able to share common key materials, and pool for each cluster has no common instance with another's.

Inter-cluster pre-distribution: In this offline sub-procedure, a number of nodes from each cluster are randomly chosen to be cluster-head (the number of cluster-heads should be enough to deliver an acceptable connectivity level, more details are provided in V-A1). All these cluster-heads are loaded with key materials from a large generated pool. This sub-procedure provides a basis for heads (clusters) to find common instances.

2) Pair discovery and pair-wise key establishment: Each sensor tries to detect its adjacent nodes (nodes within its communication range). Then, it tries to exchange information that determine: First, whether the adjacent node is from a similar cluster. Second, the adjacent node shares enough instants (based on the adopted underlying scheme) in order to establish a pair-wise key. Third, if the adjacent node satisfies the key agreement requirements (based on the adopted underlying scheme), then this pair of nodes compute (or allocate) a pair-wise key to be used in their secure peer to peer connection. A cluster-head may need to detect two classes of nodes: nodes from its legacy cluster, and cluster-heads from other clusters.

3) Path-Key establishment: In probabilistic key establishment protocols where there is likelihood for two nodes to establish a pair-wise key, there is always a fraction of nodes that may not be able to setup pair-wise key conferring to the key agreement requirements. Therefore, it is possible to improve the connectivity of network graph by completing the connection (edges) between nodes. In order to accomplish this, two adjacent nodes from the same cluster need to find a common node where both of them already have established a pair-wise key with it. This intermediate node cooperates with these adjacent nodes to supply them with a secure connection for key agreement among them. However, this procedure could be energy intensive which introduces an unnecessary communication overhead, as it is suggested to be ignored in [29] and only be considered for critical situations.

TABLE I
FREQUENTLY USED NOTATIONS

$\boldsymbol{N}$, and $\boldsymbol{n_i}$: Number of nodes, in the network, and in i^{th} cluster. $\boldsymbol{\acute{n}_i}$ denote the average number of neighboring nodes of a node in i^{th} cluster
$\boldsymbol{Y}$: Number of cluster-heads in the entire of the network, $\boldsymbol{y_i}$ denotes number of cluster-heads in i^{th} cluster
$\boldsymbol{C}$: Total number of clusters
$\boldsymbol{R}$, $\boldsymbol{R_i}$ and $\boldsymbol{\mathcal{R}}$: Ring of key material for, a node, a node from i^{th} cluster and for head's cluster, respectively
$\boldsymbol{S}$, $\boldsymbol{S_i}$ and $\boldsymbol{\mathcal{S}}$: Pool of key material that is generated for a node, a node from i^{th} cluster and for head's cluster, respectively
$\boldsymbol{t}$, $\boldsymbol{t_i}$ and $\boldsymbol{\mathcal{T}}$: Degree of a bivariate polynomial, for i^{th} and head's cluster, accordingly
$\boldsymbol{P_\mathcal{G}}$, $\boldsymbol{P_g^{(i)}}$ and $\boldsymbol{\mathcal{P}_g}$: Global connectivity for, entire of the network, i^{th} cluster and head's cluster
$\boldsymbol{P_l^{(i)}}$ and $\boldsymbol{\mathcal{P}_l}$: Local connectivity (the probability of existence of a link between two nodes) for i^{th} and head's cluster
$\boldsymbol{X}$, $\boldsymbol{x_i}$: Total number of captured nodes, in i^{th} cluster
$\boldsymbol{P_b}$, $\boldsymbol{P_b^{(i)}}$ and $\boldsymbol{\mathcal{P}_b}$: Probability of breaking a polynomial, after capturing of X nodes by attacker for, the network, i^{th} cluster, head's cluster
$\boldsymbol{P_f}$, $\boldsymbol{P_f^{(i)}}$ and $\boldsymbol{\mathcal{P}_f}$: Probability of cracking a link between two uncompromised nodes after capturing of X nodes by attacker, in network, i^{th} cluster and head's cluster
$\boldsymbol{\mathcal{P}_r}$: Resiliency (Robustness) of the network against X captured nodes
$\boldsymbol{RC}$, $\boldsymbol{RC^{(i)}}$ and $\boldsymbol{\mathcal{RC}}$: Resilient-Connectivity for, network, i^{th} and head's cluster accordingly

IV. PARADIGMS OF THE PROPOSED FRAMEWORK

To investigate the proficiency of the proposed framework, we develop two instantiations that are polynomial based pre-distribution schemes.

A. Basic Polynomial-based Instantiation (BPI)

In this instantiation, Blundo's key establishment scheme [6] is employed. Blundo's scheme utilizes a symmetric bivariate t-degree polynomial $f(x,y) = \sum_{i,j=0}^{t} a_{ij} x^i y^j$ which is randomly generated over a finite field $\mathbb{F}_Q$, where Q is a pre-determined prime number that is large enough. Prior to deployment, the key setup server computes a polynomial share for every sensor node using its assigned unique identity. To

establish a pairwise key between two sensor nodes A_a[1] and B_b, node A_a needs to evaluate its polynomial share $f(a,y)$[2] at node B_b, and similarly, node B_b evaluates its polynomial share $f(b,y)$ using identity of node A_a. Due to the desirable nature of symmetric polynomials, each pair of sensor nodes can compute a common pair-wise key to establish a secure communication, $K_{ab} = K_{ba} = f(a,b) = f(b,a)$.

1) Pre-distribution:

- For the i^{th} cluster, a unique symmetric bivariate t_i-degree polynomial $f_i(x,y)$[3] is randomly generated.
- The j^{th} sensor node belonging to the i^{th} cluster is loaded with a polynomial share $f_i(j,y)$ computed based on the node's unique identity and associated cluster polynomial.
- A unique symmetric bivariate t_i-degree polynomial $h(x,y)$ is randomly generated, and the k^{th} cluster-head is loaded with an additional polynomial share $h(k,y)$.

2) Pair discovery and pair-wise key establishment: Any two adjacent mobile sensor nodes $A_{a,u}$ and $B_{b,v}$ transmit their cluster identifiers, where $u, v \in i$. If they belong to the same cluster ($u = v$), then they exchange their unique identities, where $a, b \in j$ or k. Then, both sensor nodes evaluate their pre-loaded polynomial share at the partner's unique identity $f_u(a,b) = f_v(b,a)$, where the resulting value is their pair-wise key K_{ab}[4] $= K_{ba}$. Since every node belonging to each cluster stores one unique polynomial share , any pair of adjacent nodes from the same cluster may communicate directly and they do not need to establish a path-key.

B. Pool-based Polynomial Instantiation (PPI)

The idea of using multiple polynomials originated by Liu *et al.* [4], forms the basis of our second instantiation. In Liu's method, a set of multiple bivariat t-degree polynomials are randomly generated. Afterward a subset of polynomials are picked, and used to computed polynomial shares to be allocated for each sensor node. Then, the sensor nodes try to find at least a common polynomial share with other sensor nodes in order to establish a pair-wise key using the polynomial-based key pre-distribution scheme discussed in [6].

1) Pre-distribution phase:

- For the i^{th} cluster, a pool of multiple symmetric bivariate t_i-degree polynomials $S_i = \bigcup f_{i,p}(x,y)$ is randomly generated (where p is the polynomial identifier).
- For the j^{th} sensor node, a fixed size subset of polynomials $R_i = \bigcup f_{i,p}(j,y)$ is randomly chosen from the i^{th} cluster polynomial pool, where $R_j \subset S_i$.
- A pool of multiple symmetric bivariate t_i-degree polynomials $\mathcal{S} = \bigcup h_p(x,y)$ is randomly generated, and the k^{th} cluster-head is loaded with additional polynomial shares $\mathcal{R} = \bigcup h_p(k,y)$ where $\mathcal{R} \subset \mathcal{S}$.

[1] $\boldsymbol{A_{a,u}}$: Representation of node A, where a and u are the unique identity of node and cluster, respectively.

[2] $\boldsymbol{f_{u,p}(a,y)}$: A polynomial share for node $A_{a,u}$

[3] $\boldsymbol{f_{i,p}(x,y)}$: A bivariate polynomial with index number of p for i^{th} cluster.

[4] $\boldsymbol{K_{ab}}$: A pair-wise key for nodes $A_{a,u}$ and $B_{b,v}$, where $u = v$.

2) Pair discovery and pair-wise key establishment: Any two adjacent mobile sensor nodes $A_{a,u}$ and $B_{b,v}$ transmit their cluster identifiers, where $u, v \in i$ or k. If they belong to the same cluster ($u = v$), then they exchange a list of their polynomial identifiers α_j (α_a and α_b), as well as their unique identities, where $a, b \in j$. Both sensor nodes examine the received list of polynomial share identifiers. If they find at least one polynomial share on p^{th} polynomial, the pair-wise key is computed as $K_{ab} = K_{ba} = f_{u,p}(a,y) = f_{v,p}(b,y)$.

3) Path-key establishment: If two adjacent sensor nodes $A_{a,u}$ and $B_{b,v}$, belonging to the i^{th} cluster could not find at least one common polynomial share from p^{th} polynomial, then, they try to find an intermediate adjacent node $C_{c,w}$ (where $c \in j$ or k, and $w \in i$) from their cluster which both of them have established a pair-wise key with him. Assuming there exist $K_{a,c}$ and $K_{c,b}$, then resulting path-key may be computed at node $C_{c,w}$ as, $K_{a,b} = H[K_{a,c}||K_{c,b}]$ (Hash value of $K_{a,c}$ and $K_{c,b}$). This Path-key is encrypted via relative pair-wise key ($K_{a,c}$ or $K_{c,b}$) and sent to $A_{a,u}$ and $B_{b,v}$, then $C_{c,w}$ removes $K_{a,b}$ from its memory. Note that, this process can be communication intensive and might introduce security risk since the actual key is computed at intermediate node. In [4], there are two suggested methods for path-key establishment that can be also applied here.

V. Analysis and Assessment

In this section, we develop analytical models to evaluate the performance of the key establishment schemes within the proposed framework. Specifically, the connectivity of the framework under different scenarios is analyzed in section V-A. Resiliency against node attacks is discussed in Section V-B. Scalability and the maximum size of a network are defined in Section V-C.

A. Connectivity

1) Global and local connectivity: Irregular connection characteristic of wireless sensor networks is an inevitable issue that impacts the performance of a key establishment scheme. Therefore, connectivity is one of the performance metrics that must be highly considered in evaluation of a key establishment scheme [24]. BPI provides full (100%) global connectivity within each cluster, therefore this section focuses on the local and global connectivity of PPI.

Connected random graph: Erdös and Renyi's random graph theory [30] describes the expected node degree that lets the network stay connected. Consequently, a relationship between a desired global connectivity $P_g^{(i)}$and the local connectivity $P_l^{(i)}$, for the given n_i is obtained from

$$P_l^{(i)} = \left(\frac{1}{n_i - 1}\right) \ln \left(\frac{n_i}{\ln \frac{1}{P_g^{(i)}}}\right). \tag{1}$$

$K-$connected graph: A graph said to be $K-$connected, if any $K-1$ nodes of network is failed (or revoked), the graph is guaranteed to be still connected. Form [31], for a

$K-$connected graph and any arbitrary node distribution, $P_g^{(i)}$ is given by,

$$P_g^{(i)} = \exp\left(\frac{-n_i \Gamma(K, n_i P_l^{(i)})}{(K-1)!}\right) \tag{2}$$

where $\Gamma(a,b)$ denotes an incomplete Gamma function, defined by $\Gamma(a,b) = (a-1)!e^{-b}\sum_{i=0}^{a-1}(b^i/i!)$ [32]. From (2), when $K=2$ we have,

$$P_l^{(i)} = \left(1 - W_{-1}\left(\frac{-e \ln P_g^{(i)}}{n_i}\right)\right) \tag{3}$$

where $W_{-1}(.)$ signifies the real-valued, non-principal branch of the Lambert function, $W(x)$ [33]. For real valued x, $e^{W(x)}W(x) = x$ has two answers of $W_0 \geq -1$ and $W_{-1} \leq -1$, denoted by principal branch and non-principal branch, respectively.

Connected graph for mobile nodes: In [31], author defined a general formula to describe the global connectivity when graph includes large number of nodes ($n_i \gg 1$) as,

$$P_g^{(i)} = e^{-n_i e^{-\mu_0}}. \tag{4}$$

In (4), μ_0 is interpreted as average degree of a node in a graph, that is defined for mobile nodes by

$$\mu_0 \approx \frac{n_i \hat{r}_0^2}{3}\left(\left(4 - 2P_p + P_p^2\right) - \frac{4}{\pi}P_p^2 \hat{r}_0 - 3\left(1 - P_p\right)\hat{r}_0^2\right) \tag{5}$$

where $\hat{r}_0^2 = \dfrac{r_0^2}{a^2} = P_l^{(i)}$ [5] and P_p states the probability that a node pauses at a given time instantaneous. (5) holds when $a \gg r_0$ and mobile nodes travels autonomously of other nodes affording to the random way point (RWP) model [34]. If one assumes no pausing node in the network ($P_p = 0$), by substituting $\hat{r}_0^2$ for $P_l^{(i)}$ in (5), we have,

$$\mu_0 \approx \frac{-n_i P_l^{(i)}}{3}\left(4 - 3P_l^{(i)}\right). \tag{6}$$

Thus, from (4) and (5) we drive the value of $P_l^{(i)}$ for a desired value of $P_g^{(i)}$, when nodes are mobile, as

$$P_g^{(i)} = e^{-n_i e^{\frac{-n_i P_l^{(i)}}{3}\left(4 - 3P_l^{(i)}\right)}}. \tag{7}$$

Global connectivity of the proposed framework: $P_{\mathcal{G}}$ is inclined by the global connectivity of every clusters in the network. Therefore, for a network consists of $\mathcal{C}$ clusters, we have

$$P_{\mathcal{G}} = \mathcal{P}_g \prod_{i=1}^{\mathcal{C}} P_g^{(i)}. \tag{8}$$

Let us define $\acute{P}_g^{(i)} \triangleq 1 - P_g^{(i)}$ and $\acute{\mathcal{P}}_g \triangleq 1 - \mathcal{P}_g$. for small values of $\acute{P}_g^{(i)}$ and $\acute{\mathcal{P}}_g$ with a good approximation we have,

$$\begin{aligned} P_{\mathcal{G}} &= \left(1 - \acute{\mathcal{P}}_g\right)\prod_{i=1}^{\mathcal{C}} 1 - \acute{P}_g^{(i)} \\ &\approx 1 - \left(\acute{\mathcal{P}}_g + \sum_{i=1}^{\mathcal{C}} \acute{P}_g^{(i)}\right). \end{aligned} \tag{9}$$

Determining $|\boldsymbol{S}|$: Local connectivity in context of key establishment is the probability for two sensor nodes to establish a pair-wise key. The value of local connectivity is directly associated with value of S and R. Thus, we should know how much S should be given for a fixed value of R to hold a required local connectivity. $P_l^{(i)}$ in terms of $|S_i|$ (or $|\mathcal{S}|$), $|R_i|$ (or $|\mathcal{R}|$) and q [6] is given by,

$$P_l^{(i)} = 1 - \sum_{j=0}^{q_i - 1} \frac{\binom{R_i}{j}\binom{S_i - R_i}{R_i - j}}{\binom{S_i}{R_i}} \tag{10}$$

(10) does not consider the communication range of a sensor node. To resolve that, we assume the number of neighbor nodes $\acute{n}$ within communication range of a sensor node is given. Thus, we have $\acute{p}_l^{(i)} = \frac{n}{\acute{n}} P_l^{(i)}$.

Depending on different connectivity requirements (which is defined by an application based on different connectivity models), value of S_i (or $\mathcal{S}$) is varied. Thus, to satisfy a required local connectivity $P_l^{(i)}$ (or corresponding $P_g^{(i)}$), we need to replace $P_l^{(i)}$ in (1), (3), (7) with $\acute{p}_l^{(i)}$ for connected, $K-$connected and mobile connected, respectively.

B. Resiliency against the node capture attack

This section provides a comprehensive analysis on resiliency of our proposed framework against node capture attack. Resilience represents the compliment fraction of connections that adversary can compromise as a result of recovering key materials from captured nodes. We go on our analysis by answering the following questions (a,b and c):

a) For a network including a single cluster, what is P_b, the probability that a polynomial is being compromised after X number of nodes are captured?: When we assume a single cluster network for our analysis, similar to [4], for $N \leq t$, the probability $P_b = 0$. Here, we have considered more accurate value for P_b when $N > t$ than the study in [4]. Therefore, we drive,

$$P_b = \begin{cases} 0 & \text{if } X \leq t \\ 1 - \sum_{z=0}^{min(X,t)} \binom{X}{z}\left(\frac{R}{S}\right)^z \left(1 - \frac{R}{S}\right)^{X-z} & \text{Otherwise} \end{cases} \tag{11}$$

[5] The area in which the nodes are distributed is a disk of radius a. r_0 is the transmission range of a node, i.e., two nodes can establish a link if the distance between them is less than or equal to r_0.

[6] Assume, these two sensors need to find at least q common number of polynomial shares.

Substituting 1 for R and S in (11), results in similar security analysis in [6] which is, $P_b = \begin{cases} 0 & \text{if } X \leq t \\ 1 & \text{Otherwise} \end{cases}$.

b) What would be the probability of compromising a connection link P_f in a single cluster network?:

$$\begin{aligned} P_f &= P(\text{A link is broken} \mid \text{There exists a link}) \\ &= \frac{P(\text{A link is broken } \& \text{ There exists a link})}{P(\text{There exists a link})}. \end{aligned} \tag{12}$$

Thus,

$$P_f = \frac{\sum_{i=q}^{R} \frac{\binom{R}{i}\binom{S-R}{R-i}}{\binom{S}{R}} (P_b)^i}{\sum_{i=q}^{R} \frac{\binom{R}{i}\binom{S-R}{R_i}}{\binom{S}{R}}} \tag{13}$$

In order to crack a secure link, an attacker needs to extract the polynomials by achieving polynomial shares from captured nodes. Thus, the attacker must capture enough nodes to attain sufficient polynomial shares. The probability of compromising a secure link P_f is directly dependent on the probability of the discovering the polynomial P_b.

c) For the proposed framework (when we have multiple clusters), what is the $P_b^{(i)}$ for i^{th} cluster?:

$$\begin{aligned} P_b^{(i)} &= P(f_i \text{ is broken } |X) \\ &= \sum_{x_i=0}^{X} P(f_i \text{ is broken } |x_i, X).P(x_i|X) \end{aligned} \tag{14}$$

It is clear that if $X \leq t_i \Rightarrow P_b^{(i)} = 0$. On the other hand, x_i could not be greater than n_i and X, so upper bound of x_i is $\min(n_i, X)$. So, we have:

$$P_b^{(i)} = \sum_{x_i=t_i}^{min(n_i,X)} P(f_i|x_i, X).P(x_i|X) \tag{15}$$

and since $x_i \geq t_i$, so $\min(x_i, t_i) = t_i$, from (11) we have,

$$\begin{aligned} P(f_i|x_i, X) &= P(f_i|x_i) \\ &= 1 - \sum_{z=0}^{t_i} \binom{x_i}{z} \left(\frac{R_i}{S_i}\right)^z \left(1 - \frac{R_i}{S_i}\right)^{x_i - z}. \end{aligned} \tag{16}$$

On the other hand,

$$P(x_i|X) = \frac{\binom{\max(X,n_i)}{x_i}\binom{N-\max(X,n_i)}{\min(X,n_i)-x_i}}{\binom{N}{\min(X,n_i)}} \tag{17}$$

Let us define $M_i \triangleq \max(X, n_i)$ and $m_i \triangleq \min(X, n_i)$, by putting (16) and (17) into (15), we have $P_b^{(i)}$.

Then, similar to (13), we derive the probability of compromising a link in i^{th} cluster, by

$$P_f^{(i)} = \frac{\sum_{z=q}^{R_i} \frac{\binom{R_i}{z}\binom{S_i-R_i}{R_i-z}}{\binom{S_i}{R_i}} \left(P_b^{(i)}\right)^z}{\sum_{z=q}^{R_i} \frac{\binom{R_i}{z}\binom{S_i-R_i}{R_i-z}}{\binom{S_i}{R_i}}} \tag{18}$$

For head's cluster, similar to (15), (16) and (17), the probability of recovering one polynomial is given by

$$\begin{aligned} \mathcal{P}_b = &\sum_{\mathcal{X}=\mathcal{T}}^{\min(Y,X)} \frac{\binom{\alpha}{\mathcal{X}}\binom{N-\alpha}{\beta-\mathcal{X}}}{\binom{N}{\beta}} \\ &\times \left(1 - \sum_{z=0}^{\mathcal{T}} \binom{\mathcal{X}}{z} \left(\frac{\mathcal{R}}{\mathcal{S}}\right)^z \left(1 - \frac{\mathcal{R}}{\mathcal{S}}\right)^{(\mathcal{X}-z)}\right) \end{aligned} \tag{19}$$

where $\beta \triangleq \min(X, Y)$, $\alpha \triangleq \max(X, Y)$, we see

$$\mathcal{P}_f = \frac{\sum_{z=q}^{\mathcal{R}} \frac{\binom{\mathcal{R}}{z}\binom{\mathcal{S}-\mathcal{R}}{\mathcal{R}-z}}{\binom{\mathcal{S}}{\mathcal{R}}} (\mathcal{P}_b)^z}{\sum_{z=q}^{\mathcal{R}} \frac{\binom{\mathcal{R}}{z}\binom{\mathcal{S}-\mathcal{R}}{\mathcal{R}-z}}{\binom{\mathcal{S}}{\mathcal{R}}}} \tag{20}$$

And finally, probability of breaking a link in entire network is given by

$$P_f = \frac{\sum_{z=1}^{\mathcal{C}} \binom{n_z}{2} P_f^{(z)} + \frac{1}{2}\sum_{i=1}^{\mathcal{C}}\sum_{z=1, z\neq i}^{\mathcal{C}} y_i y_z \mathcal{P}_f}{\sum_{z=1}^{\mathcal{C}} \binom{n_z}{2} + \frac{1}{2}\sum_{i=1}^{\mathcal{C}}\sum_{z=1, z\neq i}^{\mathcal{C}} y_i y_z} \tag{21}$$

P_f in (21) represents the probability of a successful node capture attack when X nodes are captured. Therefore, the resiliency (robustness) of a network $\mathcal{P}_r$ is given by $\mathcal{P}_r = 1 - P_f$.

$P_f^{(i)}$ and $\mathcal{P}_f$ in (18) and (20), considered a more sophisticated key establishment than PPI where two nodes need to find q number of common polynomial shares in order to establish a pair-wise key. Therefore, to find P_f for PPI, $q = 1$ in (18) and (20).

In BPI, only one polynomial share is loaded into a sensor node memory. Thus, values for $R_i, S_i, \mathcal{R}, \mathcal{S}, q$ equals to 1 in (16), (18), (19), (20) , consequently,

$$P_f^{(i)} = P_b^{(i)} = \sum_{x_i=t_i}^{min(n_i,X)} \frac{\binom{\max(X,n_i)}{x_i}\binom{N-\max(X,n_i)}{\min(X,n_i)-x_i}}{\binom{N}{\min(X,n_i)}} \tag{22}$$

and,

$$\mathcal{P}_f = \mathcal{P}_b = \sum_{\mathcal{X}=\mathcal{T}}^{\min(Y,X)} \frac{\binom{\alpha}{\mathcal{X}}\binom{N-\alpha}{\beta-\mathcal{X}}}{\binom{N}{\beta}} \tag{23}$$

C. Scalability and the maximum network size

A key establishment scheme must support admission of a large number of sensor nodes in the network without loss of security, efficiency and flexibility. Many works were done to address the scalability of a key establishment scheme in term of efficiency and security [13], [20]–[24]. In one hand, connectivity (local and global connectivity) is one essential efficiency metrics. On the other hand, resiliency is a security metric that defines the robustness of scheme against node capture attacks. Unfortunately, these two metrics are contrary to each other [35]. Specifically, desiring a global connectivity $P_{\mathcal{G}}$ (or $P_g^{(i)}$) only is satisfied with providing a sufficient local connectivity $P_l^{(i)}$ and therefore for a fixed size of R (or $\mathcal{R}$), S_i (or $\mathcal{S}$) should be minimized as much as possible. Minimizing

TABLE II
MAXIMUM SIZE OF NETWORK

Scheme	$N_{\max}$	Remark
Blundo [6]	t	$\mathcal{C}=1$
Liu and Ning [4]	$\frac{(t+1)S}{R}$	$\mathcal{C}=1$
BPI	$\sum_{i=1}^{\mathcal{C}} t_i$	$Y \leq \mathcal{T}$
PPI	$\sum_{i=1}^{\mathcal{C}} (t_i+1)\frac{S}{R}$	$Y \leq (\mathcal{T}+1)\frac{\mathcal{S}}{\mathcal{R}}$

S_i (or $\mathcal{S}$) decreases the $\mathcal{P}_r$. After all, when the number of nodes N is large in the network, $P_l^{(i)}$ could be smaller to satisfy an anticipated $P_{\mathcal{G}}$ (or $P_g^{(i)}$), consequently, smaller S_i (or $\mathcal{S}$) would be needed, and indeed the value of $\mathcal{P}_r$ is reduced. Scalability can be seen in different perspective, where node density or the scale size of network defines the scalability. In fact, when N is increased in the static area of network then the node density (or $\frac{n_i}{n_i}$) is increased. Hence, $P_l^{(i)}$ is improved for the fixed S_i (or $\mathcal{S}$) and R (or $\mathcal{R}$). With a good intuition, when $\frac{n_i}{n_i}$ is large therefore attacker may find a better chance to monitor more communication and capture more nodes in low scale node capture attacks.

In [36], authors defined a metric that consider both the resiliency and connectivity that is called Resilient-Connectivity (RC). Therefore, for our proposed framework we have,

$$RC = \frac{\sum_{i=1}^{\mathcal{C}} n_i RC^{(i)} + Y\mathcal{RC}}{\sum_{i=1}^{\mathcal{C}} n_i + Y} \tag{24}$$

where $RC^{(i)} = P_g^{(i)} \times (1 - P_f^{(i)})$ and $\mathcal{RC} = \mathcal{P}_g \times (1 - \mathcal{P}_f)$.

Some scenarios of interest is to obtain the maximum possible security ($\mathcal{P}_r = 1$). Let us denote $N_{\max}$ the maximum network size which satisfies $\mathcal{P}_r = 1$. Table II shows the $N_{\max}$ for discussed instantiations in contrast with their adopted schemes. It clearly can be seen PPI and BPI are highly scalable when ultimate security is demanded as opposed to [4] and [6] respectively.

VI. NUMERICAL RESULTS

In this section we conduct extensive simulations to evaluate the security and performance of the proposed framework, and the network connectivity under different scenarios. We also compare the resiliency of aforesaid instantiations versus their adopted schemes.

The type and requirement of an assignment require a guaranteed global connectivity. To achieve a desired global connectivity, a local connectivity should be provided accordingly. Fig. 2 illustrates the desired global connectivity $P_g^{(i)}$ versus required local connectivity $P_l^{(i)}$ for cluster sizes of $n_i = 500$ and 3000 respectively, under three connectivity models of connected graph, $K-$connected graph and mobile. As it is inferred from graphs in Fig. 2, when n_i increases the chance of having more $P_l^{(i)}$ is better. $K-$connected graph requires higher local connectivity as opposed to a connected graph whereas in mobile model, less local connectivity is needed to satisfy a defined global connectivity. The reason is, when

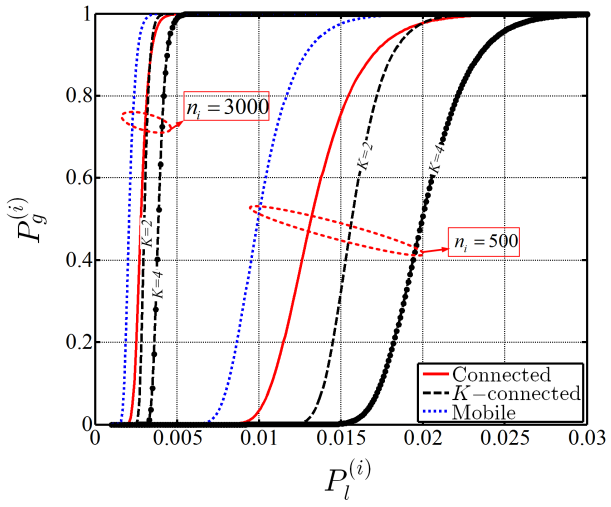


Fig. 2. Correlation of desired global connectivity $P_g^{(i)}$ and required local connectivity $P_l^{(i)}$ based on different connectivity model, for a network with n_i nodes.

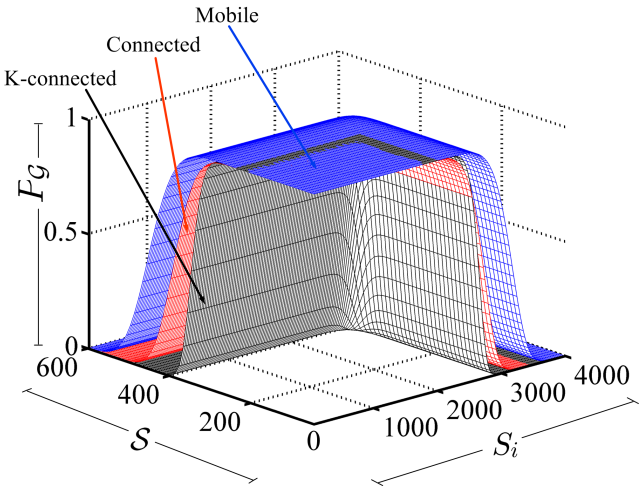
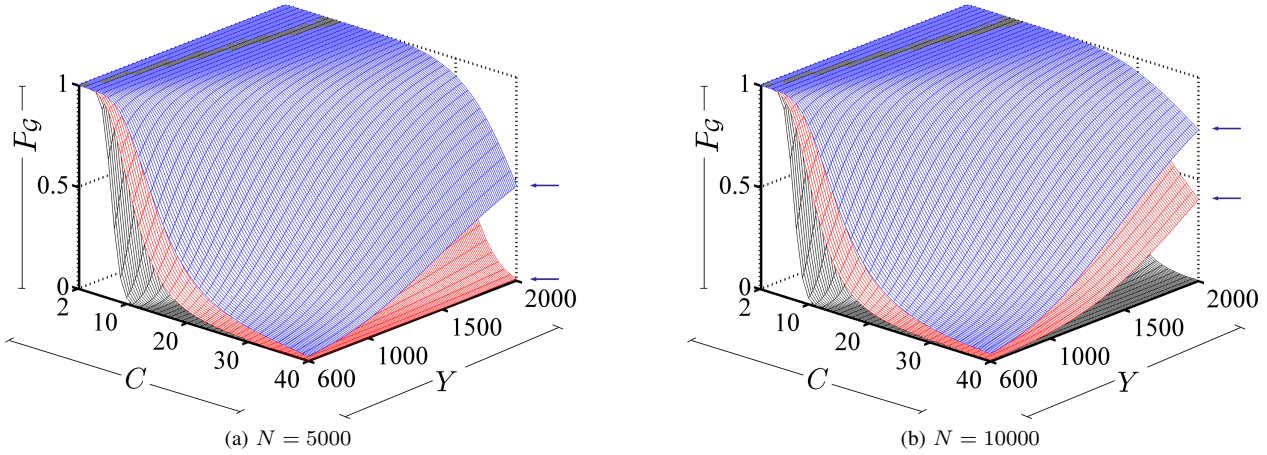


Fig. 3. Impact of $|S_i|$ and $|\mathcal{S}|$ on network global connectivity $P_{\mathcal{G}}$, for $n_i = 1000$, $y_i = 50$, $\mathcal{C} = 5$, $R_i = 20$, $\mathcal{R} = 10$ and $K = 2$ (for $K-$connected)

in mobile network an isolated node needs to communicate, it must wait until it moves or takes place into another node's radio range where it comes with a considerable communication delays. On the other hand, a high needed local connectivity $P_l^{(i)}$ may result in a resistance against movement of a node. Therefore an optimum value for local connectivity should be considered [31] which is less than $P_l^{(i)}$ for both the connected and $K-$connected graph models. Works in [37], [38] studied the optimum value for $P_l^{(i)}$ based on different assumptions.

One can see that, the maximum $|R_i|$ (or $|\mathcal{R}|$) is restricted by the hardware platform of a sensor node. Despite the fact that $|S_i|$ and $|\mathcal{S}|$ are not limited by the hardware constraints, the satisfactory value should be set for them to deliver an adequate level of connectivity. Fig. 3 makes evident $P_{\mathcal{G}}$ versus $|S_i|$ and $|\mathcal{S}|$ for different connectivity models. $P_{\mathcal{G}}$ will be decreased if a large $|S_i|$ and $|\mathcal{S}|$ is chosen. From Fig. 3, first, the impact



(a) $N = 5000$ (b) $N = 10000$

Fig. 4. Effects of increasing number of cluster $\mathcal{C}$ and cluster-head Y on the $P_{\mathcal{G}}$.

of $|\mathcal{S}|$ is more than $|S_i|$, due to the inter-connectivity that is provided by cluster-heads. In other words, a large $|\mathcal{S}|$ reduces the $\mathcal{P}_g$ and consequently $P_{\mathcal{G}}$ will be decreased. Second, mobile connectivity level is higher, and less sensitive to $|S_i|$ and $|\mathcal{S}|$ changes than both the other connectivity models.

Fig. 4 demonstrates the changes of $P_{\mathcal{G}}$ under various connectivity scenarios as the function of $\mathcal{C}$ and Y. As it is seen $P_{\mathcal{G}}$ is declined when the number of clusters increases; and is enhanced when Y is increasing. A Comparison between graphs in Fig. 4 shows the increase of N, takes less from reduction in network connectivity level.

Fig. 5 shows how the proposed framework can improve the resiliency of a network against node capture attacks. P_f decreases significantly with the increase of $\mathcal{C}$, while there is only 200 cluster-heads are employed. Increasing the cluster-heads does not improve security, as we tried the same experiment with more cluster-heads ranging from 300 to 400. Having cluster-heads imposes a nominal storage overhead. Which is more acceptable than the computation overheads as a consequences of increasing t especially at [6]. Fig. 5(a) verifies that security of [6] can be improved by suffering a bit of memory overhead (while keeping connectivity for almost 100%). Similar inference is applied to Liu and Ning scheme [4] after employing the proposed framework, as it is seen at Fig. 5(b). Additionally, comparing graphs at Fig. 5, PPI is far securer than BPI. This comes from security improvement characteristic of [4].

In order to increase $N_{\max}$, involving security parameters in Table II can be modified accordingly. $N_{\max}$ for [6] cannot be increased unusually when it is only reliant on t. t (similarly $\mathcal{T}$) cannot be a very large number as it imposes an unwanted computation overhead for a sensor node. Also, increasing of $N_{\max}$ in [4] is limited by increasing S and t, and decreasing R. Besides, Values for S and R (or R_i, $\mathcal{R}$, S_i and $\mathcal{R}$) cannot be any arbitrary values as they can aggravate the $P_{\mathcal{G}}$. For instance, an observation on Fig. 3 says that S_i and S should be carefully chosen to avoid any degradation for $P_{\mathcal{G}}$. To solve the above issue, our proposed framework makes it possible to increase $N_{\max}$ by increasing value of $\mathcal{C}$ and without need of modifying other security parameters.

VII. Conclusion

In this paper, we propose a cluster-based pre-distribution framework for key establishment schemes in mobile wireless sensor networks that consist of homogeneous sensor nodes. We actualized the feasibility of the proposed framework by adopting two pre-distribution schemes. Analysis and simulation results show that the proposed framework improves the security of adopted schemes in term of resiliency against node capture attack. Computational results have shown that, the resilience of a network can be improved effectively even with a small number of clusters, and a good connectivity of the network can be maintained at the same time with an enough number of cluster-heads. Our work has better scalability compared with its underlying adopted schemes for a given resiliency connectivity threshold, and is flexible to include mobile sensor nodes.

References

[1] X. Chen, K. Makki, K. Yen, and N. Pissinou, "Sensor network security: A survey," *Communications Surveys & Tutorials, IEEE*, 2009.

[2] L. Eschenauer and V. D. Gligor, "A key-management scheme for distributed sensor networks," in *Proceedings of the 9th ACM conference on Computer and communications security*, ser. CCS '02. ACM, 2002.

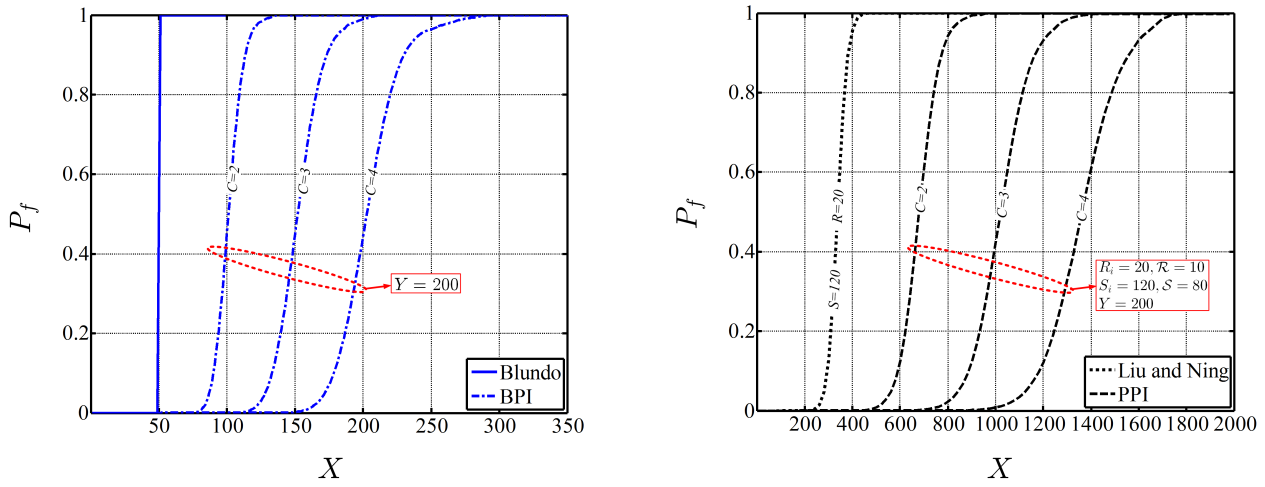
[3] H. Chan, A. Perrig, and D. Song, "Random key predistribution schemes for sensor networks," in *Security and Privacy, 2003. Proceedings. 2003 Symposium on*, 2003.

[4] D. Liu, P. Ning, and R. Li, "Establishing pairwise keys in distributed sensor networks," *ACM Trans. Inf. Syst. Secur.*, vol. 8, 2005.

[5] R. Blom, "An optimal class of symmetric key generation systems," in *Proc. of the EUROCRYPT 84 workshop on Advances in cryptology: theory and application of cryptographic techniques*. Springer-Verlag New York, Inc., 1985.

[6] C. Blundo, A. D. Santis, A. Herzberg, S. Kutten, U. Vaccaro, and M. Yung, "Perfectly-secure key distribution for dynamic conferences," 1995.

[7] W. Bechkit, Y. Challal, A. Bouabdallah *et al.*, "A new scalable key pre-distribution scheme for wsn," in *International Conference on Computer Communication Networks*, 2012.



(a) (b)

Fig. 5. Effects of increasing number of cluster $\mathcal{C}$ on the probability of P_f after X node captures, when $N = 2000$.

[8] W. Zhang, M. Tran, S. Zhu, and G. Cao, "A random perturbation-based scheme for pairwise key establishment in sensor networks," in *MobiHoc*, 2007.

[9] W. Zhang, N. Subramanian, and G. Wang, "Lightweight and compromise-resilient message authentication in sensor networks," in *INFOCOM*, 2008, pp. 1418–1426.

[10] N. Subramanian, C. Yang, and W. Zhang, "Securing distributed data storage and retrieval in sensor networks," *Pervasive Mob. Comput.*, 2007.

[11] M. Albrecht, C. Gentry, S. Halevi, and J. Katz, "Attacking cryptographic schemes based on "perturbation polynomials"," Cryptology ePrint Archive, Report 2009/098, 2009.

[12] D. Liu and P. Ning, "Improving key predistribution with deployment knowledge in static sensor networks," *ACM Trans. Sen. Netw.*, 2005.

[13] E. Klaoudatou, E. Konstantinou, G. Kambourakis, and S. Gritzalis, "A survey on cluster-based group key agreement protocols for wsns," *Communications Surveys Tutorials, IEEE*, 2011.

[14] W. Du, J. Deng, Y. Han, S. Chen, and P. Varshney, "A key management scheme for wireless sensor networks using deployment knowledge," in *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, 2004.

[15] N. T. Canh, Y.-K. Lee, and S. Lee, "Hgkm: A group-based key management scheme for sensor networks using deployment knowledge," in *Proceedings of the Communication Networks and Services Research Conference*, ser. CNSR '08. IEEE Computer Society, 2008.

[16] N. T. Canh, P. T. H. Truc, T. H. Hai, L. X. Hung, Y.-K. Lee, and S. Lee, "Enhanced group-based key management scheme for wireless sensor networks using deployment knowledge," in *Proceedings of the 6th IEEE Conference on Consumer Communications and Networking Conference*, ser. CCNC'09. IEEE Press, 2009.

[17] D. Liu, P. Ning, and W. Du, "Group-based key predistribution for wireless sensor networks," *ACM Trans. Sen. Netw.*, 2008.

[18] K. M. Martin, M. B. Paterson, and D. R. Stinson, "Key predistribution for homogeneous wireless sensor networks with group deployment of nodes," *ACM Trans. Sen. Netw.*, 2010.

[19] Z. Yu and Y. Guan, "A key management scheme using deployment knowledge for wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, 2008.

[20] Y. Xiao, V. K. Rayi, B. Sun, X. Du, F. Hu, and M. Galloway, "A survey of key management schemes in wireless sensor networks," *Comput. Commun.*, 2007.

[21] J. Zhang and V. Varadharajan, "Review: Wireless sensor network key management survey and taxonomy," *J. Netw. Comput. Appl.*, vol. 33, 2010.

[22] S. A. Camtepe and B. Yener, "Key distribution mechanisms for wireless sensor networks: a survey," Tech. Rep., 2005.

[23] Y. Wang, G. Attebury, and B. Ramamurthy, "A survey of security issues in wireless sensor networks," *Communications Surveys Tutorials, IEEE*, 2006.

[24] M. A. Simplício, Jr., P. S. L. M. Barreto, C. B. Margi, and T. C. M. B. Carvalho, "A survey on key management mechanisms for distributed wireless sensor networks," *Comput. Netw.*, 2010.

[25] K. Lu, Y. Qian, M. Guizani, and H.-H. Chen, "A framework for a distributed key management scheme in heterogeneous wireless sensor networks," *Trans. Wireless. Comm.*, 2008.

[26] M. Younis, K. Ghumman, and M. Eltoweissy, "Location-aware combinatorial key management scheme for clustered sensor networks," *Parallel and Distributed Systems, IEEE Transactions on*, 2006.

[27] M. Eltoweissy, M. Moharrum, and R. Mukkamala, "Dynamic key management in sensor networks," *Communications Magazine, IEEE*, 2006.

[28] J. Zhang and V. Varadharajan, "A new security scheme for wireless sensor networks," in *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*, 2008.

[29] F. Delgosha and F. Fekri, "A multivariate key-establishment scheme for wireless sensor networks," *Wireless Communications, IEEE Transactions on*, 2009.

[30] P. Erdös and A. Renyi, "On random graphs," in *Publication of the mathematical institute of the Hungarian academy of sciences*, 1959, pp. 290–297.

[31] C. Bettstetter, "On the connectivity of ad hoc networks," *The Computer Journal*, 2004.

[32] M. Abramowitz and I. A. Stegun, "Handbook of mathematical functions: With formulas, graphs, and mathematical tables," Dover Publications, 1965.

[33] R. Corless, G. Gonnet, D. Hare, D. Jeffrey, and D. Knuth, "On the lambert w function," 1996.

[34] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing*. Kluwer Academic Publishers, 1996.

[35] K. Chan and F. Fekri, "A resiliency-connectivity metric in wireless sensor networks with key predistribution schemes and node compromise attacks," *Physical Communication*, 2008.

[36] W. Gu, S. Chellappan, X. Bai, and H. Wang, "Scaling laws of key predistribution protocols in wireless sensor networks," *IEEE Transactions on Information Forensics and Security*, 2011.

[37] M. Zavlanos, M. Egerstedt, and G. Pappas, "Graph-theoretic connectivity control of mobile robot networks," *Proceedings of the IEEE*, 2011.

[38] A. Friend, V. Manshadi, and A. Saberi, "Distributed node placement algorithms for constructing well-connected sensor networks," in *INFOCOM, 2012 Proceedings IEEE*, 2012.