CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE
Concurrency Computat.: Pract. Exper. 2009; **21**:1373–1387
Published online 16 April 2009 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/cpe.1429

A scalable key pre-distribution mechanism for large-scale wireless sensor networks

An-Ni Shen[1], Song Guo[1,*,†], Hung Yu Chien[2] and Minyi Guo[3]

[1]*School of Computer Science and Engineering, University of Aizu, Japan*
[2]*Department of Information Management, National Chi Nan University, Taiwan*
[3]*Department of Computer Science and Engineering,*
Shanghai Jiao Tong University, China

SUMMARY

Many applications of wireless sensor network (WSN) require secure data communications, especially in a hostile environment. In order to protect the sensitive data and the sensor readings, secret keys should be used to encrypt the exchanged messages between communicating nodes. Traditional asymmetric key cryptosystems are infeasible in WSNs due to the extremely low capacity and constrained resources at each senor node. Recently proposed protocols are either vulnerable to the large-scale node capture attacks or lack of performance scalability in terms of storage, communication and computation costs. To address these limitations, we study the key pre-distribution schemes in this paper and propose a new one with all of the following properties which are particularly beneficial to the large-scale resource-constrained WSNs: (1) it completely defends against the node capture attacks, (2) it provides full connectivity of the network, and (3) it reduces the storage and communication overhead significantly compared with the other proposals. Copyright © 2009 John Wiley & Sons, Ltd.

Received 2 December 2008; Revised 30 January 2009; Accepted 12 February 2009

KEY WORDS: key management; pair-wise key; key agreement; hierarchical wireless sensor network; security

[*] Correspondence to: Song Guo, School of Computer Science and Engineering, University of Aizu, Japan.
[†] E-mail: sguo@u-aizu.ac.jp

Contract/grant sponsor: National High-Tech Research and Development Plan of China; contract/grant numbers: 2008AA01Z106, 2006AA01Z202
Contract/grant sponsor: National Natural Science Foundation of China; contract/grant numbers: 60811130528, 60725208, 60533040
Contract/grant sponsor: Shanghai Pujiang Plan; contract/grant number: 07pj14049

1. INTRODUCTION

Recent advances in micro-electro-mechanical systems (MEMS) have enabled the development of lowcost low-power wireless sensor networks (WSNs) in a wide range of applications, such as surveillance and management of battlegrounds in the military domain and monitoring of environmental, habitat, road traffic, and healthcare conditions in the civilian domain. Without relying on any predeployed network architecture [1], sensor nodes (SNs) collaborate with sense events or phenomena of interest, process the sensor data, and relay them to the consumers of the information (data collectors). Compared with other wireless or wire-line networks, WSNs have a number of unique features, e.g. stringent power, computational/memory capacity limitations, potentially large number of deployed nodes, and vulnerability to malicious attacks. These bring upon additional challenges to researchers to design efficient and scalable security protocols.

Because SNs are not tamper-proof devices, secure communications must be supported in WSNs, especially when they are deployed to a hostile environment or an unattached area. An adversary might easily capture the sensor devices to compromise their sensitive data and keys. In order to conquer such *node capture attack* problem, it is desirable to design protocols to support secure and robust pair-wise communication among any pair of sensors. At the same time, the unique features of WSNs mentioned above should also been considered in the protocol design. For example, the conventional asymmetric key cryptosystem, such as RSA [2] and Diffie-Hellman [3], cannot be implemented in SNs due to their very limited capacities.

Recent years have seen the increased attention of research concerned with the establishment of a secure channel between SNs. The first naive solution is to let all sensor devices pre-loaded the same master key. After deployment, any two nodes can use this master key to secure the communications. This approach provides full connectivity and only one key is stored in the memory of SN. However, if one SN is physically captured by an adversary, it would compromise the entire network secrecy. Another possible approach is assigning a distinct pairwise key for each pair of SNs before they are deployed. Each SN needs to store $(n-1)$ keys, where n is the size of the network. The solution provided secure against the node captured attack. Nevertheless, the scheme lacked the specific support of large network size, due to the increasing size of network, which led to the requirements for storage cost of SN also to increase linearly. Moreover, addition of new sensors to an existing sensor network is extremely difficult.

In summary, the security and efficiency requirements in a WSN may include secrecy and authentication, robustness against node capture attack, dynamic membership management (including nodes revocation and nodes addition), strong network connectivity, scalability to large-scale networks, and low complexities on memory, computation and communication overhead. These challenges motivate us to propose scalable and robust pair-wise key distribution mechanisms between sensor devices in large-scale WSNs. In particular, our methods possess the following features that are particularly beneficial to the resource-constrained WSNs: (1) defending against node captured attack, (2) providing full connectivity of the network, and (3) reducing the network storage and communication overhead significantly. Compared with the existing key distribution schemes, our security analysis and performance evaluation illustrate that our proposal has better performance than existing protocols, in terms of robustness to node captured attack, communication overhead, key storage overhead, and energy consumption.

The remainder of this paper is organized as follows. Section 2 gives an overview of related work. Section 3 presents our system model. Section 4 describes a group of protocols for our key distribution mechanism. Section 5 evaluates the security and performance of our proposal. Section 6 summarizes our findings.

2. RELATED WORK

WSNs can be broadly classified into flat WSNs and hierarchical WSNs. In a flat WSN, all senor nodes have the same computational and communication capacities. In a hierarchical WSN, however, some special sensor devices, called Cluster Head (CH), have much higher capacities than other SNs. By applying some clustering algorithms like [4], the whole set of sensor devices could be partitioned into several distinct clusters such that each cluster has at least one CH. Under this arrangement, each SN forwards the generated packets to its local CH by short-range transmissions, and the CH then performs a pre-processing for the raw data received from all other senor nodes in the cluster and finally forwards the aggregated data to the sink node, or BS, directly by long-range transmissions.

2.1. Key distribution protocols in flat WSNs

Eschenauer and Gligor introduced the idea of probabilistic key sharing by the first time and proposed a basic random key pre-distribution scheme [5] for flat WSNs. Before deployment, each SN is assigned a subset of keys randomly chosen from a large key pool. Each SN broadcasts the indexes of their keys and finds one common key in their key sets to be used as their shared secret key. This scheme can achieve high connectivity with only moderate storage overhead. Based on the key pool and random key pre-distribution, Lin and Ning proposed the polynomial-based key pre-distribution scheme [6], which exploits the t-degree property of symmetric polynomial function [7] to improve network security. Pietro *et al.* proposed a pseudo-random key pre-distribution scheme [8], in which each senor node stores the pre-assigned keys generated by a pseudo-random generator [9] with the Id of the SN as the seed. This feature supports a key discovery phase with no message exchange required such that energy consumption can be reduced. Du *et al.* proposed a multi-space key agreement scheme [10] based on Blom's key agreement [11]. The salient characteristic of Blom's scheme, known as λ-secure property, is that as long as no more λ nodes are compromised, the networks are perfectly secure. However, recent research has shown that the schemes based on random key pre-distribution [5,6,8,10,12] and grid-based key pre-distribution [6,13] are vulnerable to node captured attacks and can only provide probabilistic connectivity of the network in; some SNs may be isolated in the network, due to the nodes not storing the common secrecy information to establish communication key with its neighbor member Cheng and Agrawal partially solved those problems by proposing the EPKEM scheme [14], which guarantees the full connectivity between sensor devices. However, the EPKEM scheme was unable to resist the node captured attack [15].

2.2. Key distribution protocols in hierarchical WSNs

Recent research on key distribution protocols in the literature has focused more on the hierarchical architecture for large-scale resource-constrained WSNs, because it has been shown in [16] that a

DOI: 10.1002/cpe

hierarchical architecture can provide a better performance, in terms of communication overhead than a flat architecture in such networks.

To solve the key agreement problem in hierarchical WSNs, Gaurave *et al.* proposed a key pre-distribution scheme Low-Energy Key Management (LEKM) [17]. Before deployment, each CH stores a set of keys in its memory and each SN randomly selects a key from a CH and stores it with the CH's Id in its memory. After deployment, each SN establishes a secure link with the CH that has been selected. This is done at each SN by exchanging key information with its CH if the CH has the key in its memory. Otherwise, the CH requests the intended key from the corresponding CH. Once CHs establish key with each of its member nodes, all pre-assigned keys are erased. Such scheme has no computational cost at both SN and CH in key establishment phase and is robust against node capture attack after the key establishment phase. On the other hand, it has high storage and communication overhead at CHs.

As another proposal for hierarchical WSNs, a polynomial-based protocol Improved Key Distribution Mechanism (IKDM) is proposed in [16]. In the IKDM scheme, each SN and CH both have fixed storage cost in pre-distribution phase and also provided full connectivity of the network. In order to improve the resilience against node captured attack, the pre-loading secrecy key of each SN is the exclusive-or a result of ℓ number of ($\ell \geq 1$) bivariate polynomial keys which can be fetched by its CH from ℓ number of distinctive CHs all over the network. The parameter ℓ defines the tradeoff between the communication overhead and the robustness to the node capture attacks at the CHs. While the large ℓ can improve the security level of the network, it will also result in significant message exchanges at the same time for establishing secure links.

3. NETWORK MODEL

In this paper, we consider a three-tier hierarchical architecture of a large-scale WSN. As illustrated in Figure 1, our model has three different types of wireless sensor devices: BS, CH and normal SN. Each low-cost SN has low data processing capability, limited memory storage and battery power supplies, and short radio transmission range. SNs are restricted to direct communications with its CH only. The CHs are equipped with high power batteries, large memory storages, powerful antenna and data processing capacities, and thus can execute relatively complicated numerical operations. It also has much longer radio transmission range than SNs such that CHs can communicate with each other in a multi-hop fashion. As the most powerful node in a WSN the BS has virtually unlimited computational and communication power, unlimited memory storage capacity, and very large radio transmission range. While both SN and CH keep stationary during the network operation period after deployment, the BS could be either stationary or mobile to ensure the full coverage of the whole network.

We apply a two-phase node deployment process to practically construct such hierarchical networks. Initially, a larger number of SNs are arbitrarily deployed in a given area. Some clustering algorithm, e.g. [4] is thereafter used to partition the whole network into several distinct clusters. The second deployment process is to place one CH at an appropriate location within each resulting cluster such that the CH can communicate with all SNs in the cluster directly.

A three-tier hierarchical WSN can thus be modeled as a simple graph G with a finite node set, including a BS, m CHs and n SNs. For convenience, we assume that each cluster has one CH

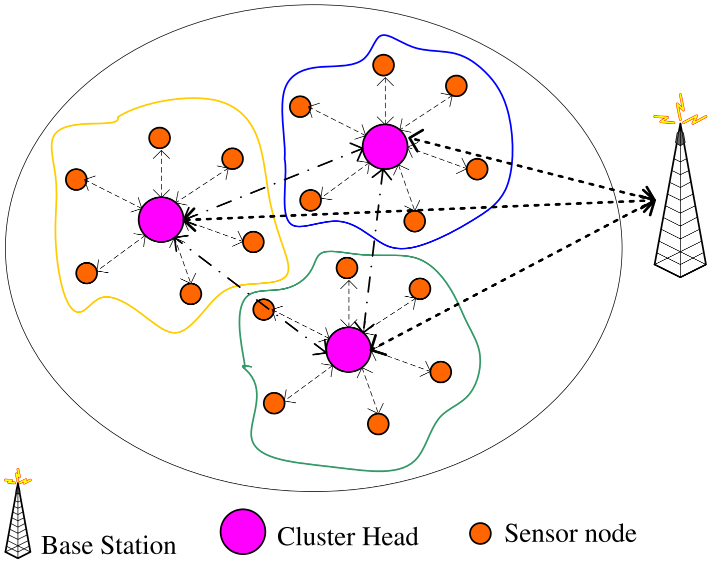

Figure 1. A three-tier hierarchical WSN.

and $\lceil n/m \rceil$ sensors inside. A secure wireless link corresponding to the wireless communication channel belongs to the arc set of G only if there exists a pair-wise key between the transmission nodes of the link. In our key pre-distribution scheme, a bivariate symmetric polynomial function is used to generate the key for each link of the network. The t-degree bivariate symmetric polynomial function $f(x, y)$, introduced in [7], is defined as follows:

$$f(x, y) = \sum_{i,j=0}^{t} a_{ij} x^i y^j \tag{1}$$

The coefficients a_{ij} $(0 \leq i, j \leq t)$ are randomly chosen from a finite field $GF(Q)$, in which Q is a prime number that is large enough to accommodate a cryptographic key. As implied by its name, the symmetric property of a bivariate polynomial function given in (1) satisfies

$$f(x, y) = f(y, x) \tag{2}$$

For any node a with a unique identifier Id_a, the preloaded polynomial $f_a(y)$:

$$f_a(y) \equiv f(Id_a, y) = \sum_{i=0}^{t} A_i(Id_a) \cdot y^i \tag{3}$$

is stored in its memory in the form of a set of symmetric coefficients $A_i(Id_a)$, i.e.

$$A_i(Id_a) = \sum_{j=0}^{t} a_{ij} \cdot (Id_a)^j \tag{4}$$

DOI: 10.1002/cpe

We assume that the polynomials $f_u(y)=f(Id_u, y)$ and $f_v(y)=f(Id_v, y)$ are preloaded at nodes u and v, respectively. After substituting the Id of the other party into variable y of the preloaded polynomial, the generated keys $f_u(Id_v)=f(Id_u, Id_v)$ and $f_v(Id_u)=f(Id_v, Id_u)$ at nodes u and v, respectively, are the same due to the symmetric property shown in (2). By applying the symmetric property, a secure link can be easily built up by just exchanging the Ids of transmission nodes. On the other hand, a t-degree bivariate polynomial key scheme can only keep secure against coalitions of up to t compromised sensors. When the number of compromised nodes is less than t, the coefficients of the polynomial cannot be derived even all the compromised nodes put their stored information together. But once more than t nodes are compromised, the adversary can crack the coefficients of the polynomial such that all the pair-wise keys in the entire group would be cracked. Although increasing the value of t can improve the security property of the bivariate polynomial key scheme, it is not suitable for WSNs due to the limited memory size of the sensors.

4. A NEW KEY PRE-DISTRIBUTION MECHANISM

To address the limitations of current key pre-distribution schemes, e.g. [16,17], we shall propose a Scalable Key Pre-distribution (SKPD) mechanism for large-scale WSNs in this section. Based on a three-tier hierarchal network model as described in Figure 1 and the polynomial key calculation mechanism, our scheme enables any pair of communicating parties (CH to sensor, CH to CH, CH to BS) to establish a unique pairwise key between them. The proposed SKPD mechanism can provide sufficient security for large-scale sensor networks against node capture attack. Furthermore, each node has zero communication overhead except for the local Id exchanges and fixed key storage overhead regardless of the network size and density.

Our approach has four phases: key pre-assignment phase, inter-cluster pairwise key establishment phase intra-cluster pairwise key establishment phase and obsolete parameter erasure phase. In order to present SKPD in a formal manner, we list the notations used in our protocol descriptions in Table I for convenience to the reader.

4.1. Key pre-assignment phase

Owing to the resource constraints of wireless sensors, some secret information needs to be preloaded into SNs and CHs before they are deployed into usage in our proposed SKPD mechanism.

Table I. Notations.

Notation	Description
BS	A base station
CH_i	Cluster head i with Id Id_{CH_i}
SN_i	Sensor node i with Id Id_{SN_i}
$K_{a,b}$	A pair-wise key between a and b
$E(\text{data}, K_{a,b})$	A symmetric encryption function using $K_{a,b}$ as a key
$f_{CH-CH}(x, y)$	A symmetric polynomial function ($s.p.f.$) used to establish the key between CHs
$f_{CH-SN}(x, y)$	A $s.p.f.$ used to establish the key between a CH and an SN
$H(x)$	A one-way hash function

DOI: 10.1002/cpe

This secrecy information to be assigned into the three types of network devices described as follows.

1. The BS has $n+m$ pair-wise keys to share with SNs and CHs respectively, for their authenticated communications. Each key is shared with a particular SN or CH. We use K_{BS,CH_i}, $1 \leq i \leq m$, to represent the shared pairwise key between BS and cluster head CH_i and K_{BS,SN_i}, $1 \leq i \leq n$, to represent the shared pairwise key between BS and sensor node SN_i.
2. Each CH_i stores a pair-wise key K_{BS,CH_i} and two symmetric polynomials $f_{CH-CH}(Id_{CH_i}, y)$ and $f_{CH-SN}(Id_{CH_i}, y)$, in which the symbol y denotes the Id of a neighboring CH CH_i $(1 \leq i \leq m)$ and an SN, SN_i $(1 \leq i \leq n)$ in its cluster, respectively. The key K_{BS,CH_i} is used to authenticate and secure the communication between BS and cluster head CH_i. Note that each polynomial is stored in memory as $(t+1)$ number of coefficients of $f_{CH-CH}(Id_{CH_i}, y)$ or $f_{CH-SN}(Id_{CH_i}, y)$.
3. Each search node SN_i, is pre-loaded a pair-wise key K_{BS,SN_i} and a symmetric polynomial $f_{CH-SN}(Id_{SN_i}, y)$. The key K_{BS,SN_i} is used to authenticate and secure the communication between BS and secure node SN_i.

In particular, we would like to point out that the space requirements to the network devices described above are independent of the size of the network. After the key pre-assignment phase, wireless sensors are randomly distributed in a given area, and later on, some clustering algorithms, e.g. [4], shall organize the network into a hierarchical structure.

4.2. Inter-cluster pair-wise key establishment phase

After the node deployment, each CH needs first to establish pairwise keys with other CHs, which are within its transmission range to secure the communication between them. Suppose CH_a and CH_b are going to establish a pair-wise key with each other. They first broadcast their own Ids and then independently calculate their keys as follows, respectively:

$$CH_a:\ K_{CH_a,CH_b} = H(f_{CH-CH}(Id_{CH_a}, Id_{CH_b})) \tag{5}$$

$$CH_b:\ K_{CH_b,CH_a} = H(f_{CH-CH}(Id_{CH_b}, Id_{CH_a})) \tag{6}$$

Note that the symmetric property ensures the same key to be obtained at both sides and the one-way hash function [18,19] makes the resulting shared key even harder to be cracked.

4.3. Intra-cluster pair-wise key establishment phase

After the inter-cluster pair-wise key establishment phase, each CH need to establish pair-wise keys with its cluster members to secure the intra-cluster communications. Assume that SN_i is a member of cluster CH_a. Similar to the previous phase, the local Id broadcasting let CH_a and SN_i know each other's Id and then they calculate their shared unique pair-wise key as follows:

$$SN_i:\ K_{CH_a,SN_i} = H(f_{CH-SN}(Id_{SN_i}, Id_{CH_a})) \tag{7}$$

$$CH_a:\ K_{SN_i,CH_a} = H(f_{CH-SN}(Id_{CH_a}, Id_{SN_i})) \tag{8}$$

DOI: 10.1002/cpe

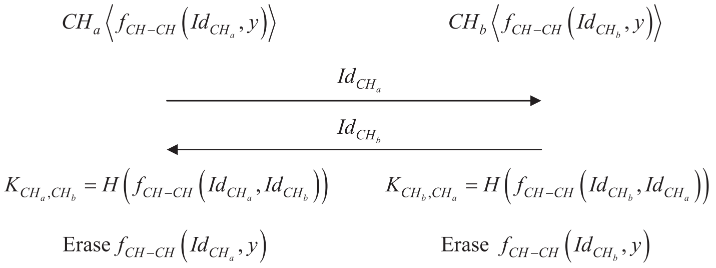
$CH_a\left\langle f_{CH-CH}\left(Id_{CH_a},y\right)\right\rangle$
$CH_b\left\langle f_{CH-CH}\left(Id_{CH_b},y\right)\right\rangle$
Id_{CH_a}
Id_{CH_b}
$K_{CH_a,CH_b}=H\left(f_{CH-CH}\left(Id_{CH_a},Id_{CH_b}\right)\right)$
$K_{CH_b,CH_a}=H\left(f_{CH-CH}\left(Id_{CH_b},Id_{CH_a}\right)\right)$
Erase $f_{CH-CH}\left(Id_{CH_a},y\right)$
Erase $f_{CH-CH}\left(Id_{CH_b},y\right)$

Figure 2. Inter-cluster key establishment.

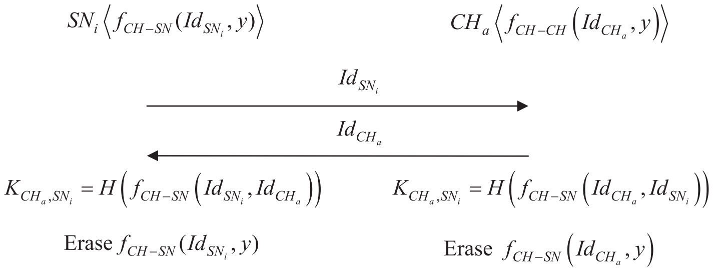
$SN_i\left\langle f_{CH-SN}(Id_{SN_i},y)\right\rangle$
$CH_a\left\langle f_{CH-CH}\left(Id_{CH_a},y\right)\right\rangle$
Id_{SN_i}
Id_{CH_a}
$K_{CH_a,SN_i}=H\left(f_{CH-SN}\left(Id_{SN_i},Id_{CH_a}\right)\right)$
$K_{CH_a,SN_i}=H\left(f_{CH-SN}\left(Id_{CH_a},Id_{SN_i}\right)\right)$
Erase $f_{CH-SN}(Id_{SN_i},y)$
Erase $f_{CH-SN}\left(Id_{CH_a},y\right)$

Figure 3. Intra-cluster key establishment.

4.4. Obsolete parameter erasure

Once each SN and head cluster has computed its pair-wise keys with all its neighboring nodes, it erases the security critical parameters, i.e. the polynomial coefficients, from its memory except for the pair-wise keys just calculated. This process prevents from the attacks caused by the compromised nodes. Finally, the complete descriptions of our proposed SKPD mechanism for inter-cluster key establishment and intra-cluster key establishment are summarized and illustrated in Figures 2 and 3, respectively.

5. SECURITY ANALYSIS

As SNs are often deployed in unattached area and they are not tamper resistant due to their low cost the adversary may be able to easily capture the sensor devices to compromise their stored sensitive data and communication keys. Some SNs' capture may compromise the communication between other non-captured nodes. This is recognized as the node capture attack, which is a serious threat in WSNs. In this section, we evaluate the security property of our proposed mechanism in terms of the ability to defense against the node capture attack. Let $F_{\mathrm{CH}}^{i}(x)$ and $F_{\mathrm{SN}}^{i}(x)$ be the fractions of compromised keys in non-captured SNs as a function of the number of compromised CHs and the number of SN, respectively, by the key distribution algorithm i. We shall compare SKPD with the LEKM [17] and the IKDM [16].

In particular, we consider the security property of all these schemes in two typical scenarios: the ability to defense against the node capture attack (1) after deployment of the network and (2) in initialization of the network (i.e. the keys have not been created yet and/or some obsolete security critical parameters have not been erased yet for all the nodes). The boostup time is normally short. For example, it only takes about 0.13 s for calculating a 80 bit key in each SN in our experiments. The probability to have a large number of SNs to be compromised in so short period of time is low. Therefore, the security property in the first scenario should be the major concern in evaluating this group of key distribution schemes in the rest of the section.

5.1. Security analysis after deployment of the network

Because only pairwise keys are remaining in the SNs for all schemes after deployment of the network, i.e. all security parameters that will not be used in the future have been already erased from the network, any SN's compromising will not endanger the secret communications of other non-captured node. In other words, all these schemes have full ability to defense the node capture attack at SNs, i.e.

$$F_{\mathrm{SN}}^{\mathrm{SKPD}}(x) = F_{\mathrm{SN}}^{\mathrm{IKDM}}(x) = F_{\mathrm{SN}}^{\mathrm{LEKM}}(x) = 0 \tag{9}$$

In the following, we turn our attention to the more challenging security analysis for CHs. In SKPD, the polynomials at each CH are erased after the inter-cluster key establishment and intra-cluster key establishment. Because the remaining pairwise keys in the CHs are unique and they cannot be used to obtain the corresponding polynomial reversely, we conclude that our SKPD mechanism has the full ability to defend the node capture attack at CHs. This conclusion applies to LEKM as well because all unrelated keys are removed at CHs after network deployment. In summary, we have

$$F_{\mathrm{CH}}^{\mathrm{SKPD}}(x) = F_{\mathrm{CH}}^{\mathrm{LEKM}}(x) = 0 \tag{10}$$

Recall that the key of each SN in IKDM is composed of ℓ number of secret shares, each of which can be calculated by a distinctive CH in the network. Furthermore, all preloaded polynomials at each CH will not be removed after network deployment. Once a group of CHs (up to x) are captured, all the keys in non-captured nodes will also be compromised if their corresponding ℓ number of secret shares can be obtained from the polynomials of these x CHs. In general, F_{IKDM} of the IKDM scheme after network deployment can be expressed as follows:

$$F_{\mathrm{CH}}^{\mathrm{IKDM}}(x) = \begin{cases} 1, & x \geq t \\ \binom{x}{\ell} \Big/ \binom{m}{\ell}, & t > x \geq 0 \end{cases} \tag{11}$$

We perform a security evaluation on a typical network configuration as in [20], in which there are $n = 10\,000$ SNs and $m = 100$ CHs in a network. The 80-degree (i.e. $t = 80$) bivariate polynomial shares are pre-loaded in each CH in SKPD and IKDM. The security results of all schemes are

DOI: 10.1002/cpe

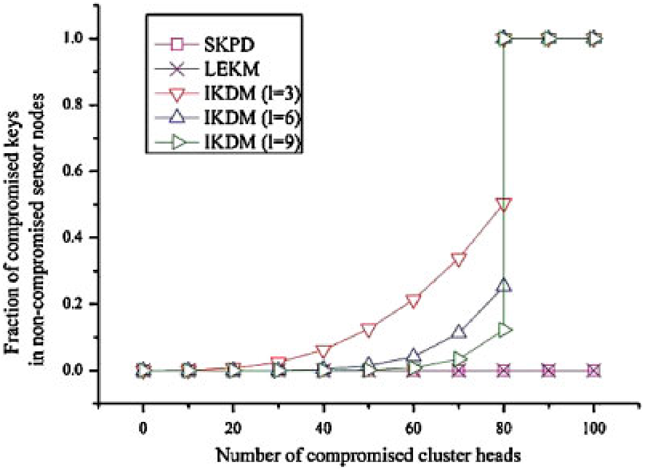

Figure 4. Fraction of compromised keys in non-captured sensor nodes vs number of compromised cluster heads after deployment of the network.

illustrated in Figure 4. We observe that both LEKM and SKPD are perfectly secure. When the number of captured CHs increases, the number of compromised sensors increases in IKDM scheme. This increasing trend is dramatic especially when ℓ is small, e.g. $\ell=3$.

5.2. Security analysis in initialization of the network

Because our scheme maintains the polynomials in the initialization, it suffers the t-security problem for both SNs and CHs, i.e.

$$F_{\text{SN}}^{\text{SKPD}}(x)=F_{\text{CH}}^{\text{SKPD}}(x)=\begin{cases}0, & 0\leq x<t\\ 1, & x\geq t\end{cases} \tag{12}$$

In the IKDM scheme, there is no such obsolete parameter erasure phase and thus it has the same security property as we discussed in the first scenario. In LEKM, SNs are perfectly secure against the node capture attack due to the fact that only pairwise keys remain at each SN. On the other hand, each CH stores n/m sensor's secret keys in its memory and therefore any single CH's capture could compromise the n/m sensors' secret keys. When the number of captured CHs increases, the number of compromised sensors increases dramatically. In summary, we have

$$F_{\text{SN}}^{\text{LEKM}}(x) = 0 \tag{13}$$

$$F_{\text{CH}}^{\text{LEKM}}(x) = \frac{x}{m}, \quad 0\leq x\leq m \tag{14}$$

The fraction of compromised keys in non-captured SNs as a function of the number of compromised SNs and the number of CHs in network initialization for all schemes are illustrated in

DOI: 10.1002/cpe

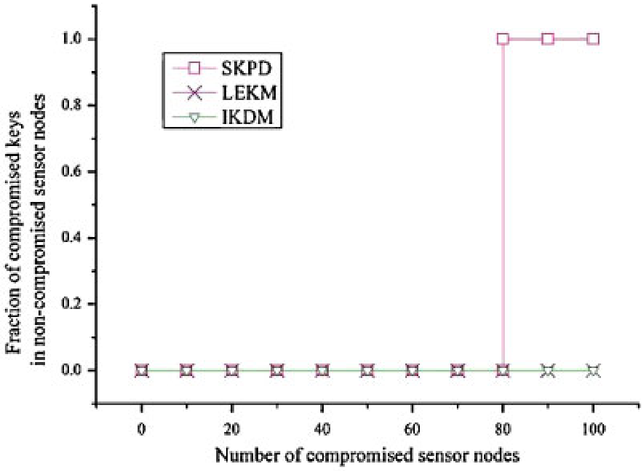

Figure 5. Fraction of compromised keys in non-captured sensor nodes vs number of compromised sensor nodes in initialization of the network.

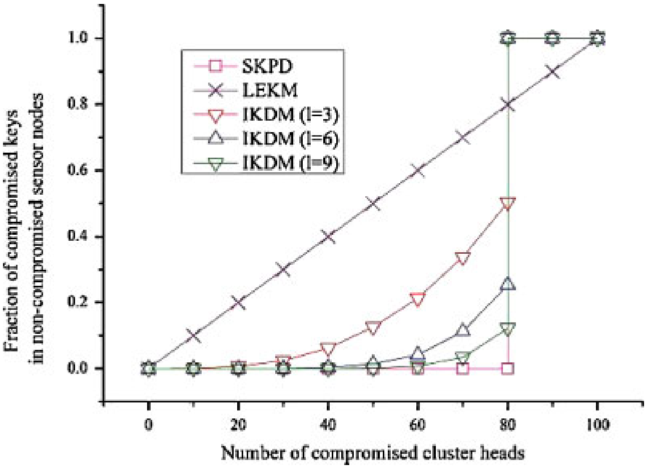

Figure 6. Fraction of compromised keys in non-captured sensor nodes vs number of compromised cluster heads in initialization of the network.

Figures 5 and 6, respectively, based on the same network configuration as the previous subsection. Our security analysis for both scenarios shows that our proposal is more robust to node captured attack than the existing protocols.

6. PERFORMANCE EVALUATION

In this section, we evaluate the performance of our proposed scheme by comparing with LEKM [17] and IKDM [16]. The performance metrics include the complexities of storage overhead, computational cost, and communication overhead.

6.1. Storage overhead

In the scheme LEKM, the storage overhead is fixed after key establishment phase. However, the number of keys stored in each CH is linearly proportional to the size of the network at pre-deployment phase. In particular, the number of keys stored in each CH as a function of n and m can be expressed as $n/m+m$. Both IKDM and our scheme have fixed storage overhead for SNs and CHs under various network sizes as shown in Table II. The above analysis shows that our scheme has better scalability in terms of storage overhead.

6.2. Computational cost

The computational cost is summarized in Table III. In our proposed scheme, the computational cost of SN is $O(t)$ using Cramer's Rule [21]. The computational cost of CH is $O(tn/m)$, where n/m denotes the average number of SNs in a cluster. IKDM scheme has no computational cost at SNs and a computational cost $O(\ell \cdot tn/m)$ at each CH, in which ℓ is the number of polynomial shares assigned to each SN chosen by an off-line key distribution server. Finally, the LEKM scheme has zero computational cost at both CHs and SNs.

The above analysis shows that our scheme has a slightly higher computational cost than other proposals. However, in a typical WSN, the radio communications have significantly higher energy consumption than the code executions and calculations [16,22]. The experimental results in [23] show that the energy of sending 1 bit to 100 m away is roughly the same as executing 3000 instructions in SNs. The extremely low communication cost, as shown in the next section, allows our scheme to achieve minimal overall energy consumption compared with other proposals.

Table II. Storage cost comparison over various distribution schemes.

	SKPD	IKDM	LEKM
Cluster head	One key Two polynomial functions	One key Two polynomial functions	$\frac{n}{m}+m$ keys
Sensor node	One key One polynomial function	Two keys ℓ Ids	One Id Two keys

Table III. Computational cost comparison over various distribution schemes.

	SKPD	IKDM	LEKM
Cluster head	$O(tn/m)$	$O(\ell \cdot tn/m)$	0
Sensor node	$O(t)$	0	0

DOI: 10.1002/cpe

6.3. Communication overhead

In order to study the scalability of communication overhead, we conduct a simulation study over network examples with different sizes of SNs from $n=90$ up to 10 000. The whole networks are regularly organized as $\sqrt{m}\times\sqrt{m}$ number of clusters ($m=9, 16, \ldots, 100$) and there are exactly 100 SNs in each $R\times R$ cluster. The transmission range of each CH is set as $\sqrt{5}R$ and the communications between CHs may be made in a multi-hop manner if they are separated far away from each other.

For key pre-distribution schemes, the communication overhead mainly occurs in the network initialization phase, since each sensor needs to exchange key information with its neighbors. In our simulation study, we defined the normalized communication overhead as the average number of bits used in communications for key establishment between an SN and its CH, in which the Id and key take up 32 and 80 bits, respectively. Note that the Id exchanges are considered as a basic and necessary operation for neighbor discovery in WSNs and thus are not counted in for the communication overhead of key distribution schemes.

In the IKDM scheme, the pre-loaded key in SN composing of ℓ secret shares, must be rebuilt in its CH by fetching from exact ℓ number of CHs. As we observed in Figure 7, the communication overhead is an increasing function of ℓ under a fixed network size. On the other hand, these ℓ number of CHs may be parted farther, in terms of transmission hops in large-scale networks due to the arbitrary deployment of the SNs. It can be observed that the normalized communication overhead is an increasing function of network size under the fixed ℓ. The performance of LEKM is also given in Figure 7 and seems lower than IKDM. This is because a CH in LEKM needs to obtain the secret share (i.e. key) from only one, instead of $\ell>1$, CH in the network. The communication cost of our proposed scheme reaches minimum because there is no extra message involved to establish keys between pairs of sensor devices.

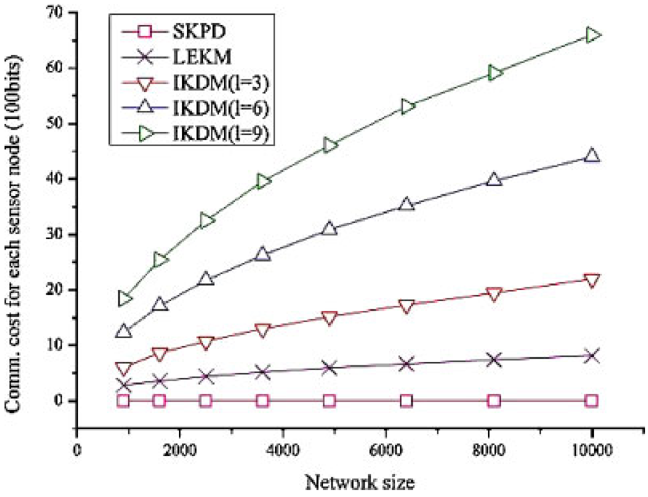

Figure 7. Communication overhead vs network size.

7. CONCLUSION

In this paper, we present an efficient and scalability key pre-distribution scheme for large-scale WSNs based on a three-tier hierarchical network architecture and bivariate polynomial-key pre-distribution mechanism. Compared with existing key distribution protocols, our scheme provides the best network resilience against node capture attack. The communication overhead of our scheme is minimized to zero, resulting in the lowest energy consumption. In our scheme, each SN only needs to store one key and one polynomial in its memory regardless of the network size and density, which reduces the key storage overhead for tiny sensors and makes our scheme suitable for large-scale WSNs.

ACKNOWLEDGEMENTS

This work is supported by the National High-Tech Research and Development Plan of China (863 Plan) under Grant Nos. 2008AA01Z106 and 2006AA01Z202, the National Natural Science Foundation of China under Grant Nos. 60811130528, 60725208, and 60533040, and Shanghai Pujiang Plan under grant No. 07pj14049.

REFERENCES

1. Akyildiz IF, Su W, Sankarasubramaniam Y, Cayirci E. Wireless sensor networks: A survey. *Journal of Computer Networks* 2002; **38**(4):393–422.
2. Rivest R, Shamir A, Adleman L. A method for obtaining digital signatures and public key cryptosystems. *Communications of ACM* 1978; **21**:120–126.
3. Diffie W, Hellman ME. New direction in cryptography. *IEEE Transactions on Information Theory* 1976; **IT-22**(6):228–258.
4. Heinzelman WR, Chandrakasan AP, Balakrishnan H. An application specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications* 2002; **1**(4):660–670.
5. Eschenauer L, Gligor V. A key-management scheme for distributed sensor networks. *ACM Conference on Computer and Communications Security*, Washington, DC, U.S.A., 2002; 41–47.
6. Liu D, Ning P. Establishing pairwise keys in distributed sensor networks. *ACM Conference on Computer and Communications Security*, Washington, DC, U.S.A., 2003; 52–61.
7. Blundo C, Santis AD, Herzberg A, Kutten S, Vaccaro U, Yung M. *Perfectly-secure Key Distribution for Dynamic Conferences* (*Lecture Notes in Computer Science*). Springer: Berlin, 1993; 471–486.
8. Pietro RD, Mancini LV, Mei A. Efficient and resilient key discovery based on pseudo-random key pre-deployment. *International Parallel and Distributed Processing Symposium*, Santa Fe, NM, U.S.A., 2004; 26–30.
9. Goldreich O, Goldwasser S, Micali S. How to construct random function. *Journal of the ACM* 1986; **33**(4):792–807.
10. Du WL, Deng J, Han Y, Varshney PK. A pairwise key pre-distribution scheme for wireless sensor network. *ACM Conference on Computer and Communications Security*, Washington, DC, U.S.A., 2003; 42–51.
11. Blom R. An optimal class of symmetric key generation system. *EUROCRYPT'84*, Paris, France (*Lecture Notes in Computer Science*, vol. 209). Springer: Berlin, 1985; 335–338.
12. Kausar F, Hussain S, Park JH, Massod A. A key distribution scheme preventing collusion attacks in ubiquitous heterogeneous sensor network. *Embedded and Ubiquious Computing Workshops* (*Lecture Notes in Computer Science*). Springer: Berlin, 2007; 745–757.
13. Sadi MG, Kim DS, Park JS. GBR: Grid based random key predistribution for wireless sensor network. *International Conference on Parallel and Distributed System*, Fuduoka, Japan, vol. 2, 2003; 310–315.
14. Cheng Y, Agrawal DP. Efficient pairwise key establishment and management in static wireless sensor networks. *Mobile Adhoc and Sensor Systems Conference*, Washington, DC, U.S.A., 2005; 544–550.
15. Chien HY, Ching RC, Shen AN. Efficient key pre-distribution for sensor nodes with strong connectivity and low storage space. *International Conference on Advanced Information Networking and Application*, Gino-wan, Okinawa, Japan, 2008; 327–333.

DOI: 10.1002/cpe

16. Cheng Y, Agrawal DP. A improved key distribution mechanism for large-scale hierarchical wireless sensor networks. *Journal of Ad Hoc Networks* 2007; **5**(1):35–48.
17. Jolly G, Kuscu MC, Kokate P, Yuonis S. A low-energy management protocol for wireless sensor networks. *IEEE International Symposium Computers and Communication*, Kemer-Antalya, Turkey, 2003; 335–340.
18. Damgard IB. A design principle for hash functions, advances. *Cryptology-CRYPTO'89*, Santa Barbara, CA, U.S.A. (*Lecture Notes in Computer Science*, vol. 435). Springer: Berlin, 1990; 416–427.
19. Merkle R. One-way hash functions and DES, advances in cryptology. *CRYPTO'89* (*Lecture Notes in Computer Science*, vol. 435). Springer: Berlin, 1989; 428–446.
20. Zhang W, Tran M, Zhu S, Cao G. A random perturbation-based scheme for pairwise key establishment in sensor networks. *ACM MobiHoc'07*, Montreal, Que., Canada, 2007; 90–99.
21. Boyer CB. *A History of Mathematics* (2nd edn). Wiley: New York, 1968.
22. Schurgers C, Tsiatis V, Ganeriwal S, Srivastava S. Optimizing sensor networks in the energy-latency-density design space. *IEEE Transactions on Mobile Computing* 2002; **1**(1):70–80.
23. Sun LM, Li JZ, Chen Y, Zhu HS. *Wireless Sensor Networks*. Tsinghua University Press: Beijing, 2005.

DOI: 10.1002/cpe