ISR TECHNICAL REPORT 2011-07

On the resiliency of sensor networks under the pairwise key distribution scheme

Osman Yagan
Armand Makowski

On the resiliency of sensor networks under the pairwise key distribution scheme

Osman Yağan and Armand M. Makowski
Department of Electrical and Computer Engineering
and the Institute for Systems Research
University of Maryland, College Park
College Park, Maryland 20742
oyagan@umd.edu, armand@isr.umd.edu

***Abstract*— We investigate the security of wireless sensor networks under the pairwise key distribution scheme of Chan et al. [2]. We present conditions on how to scale the model parameters so that the network is i) unassailable, and ii) unsplittable, both with high probability, as the number of sensor nodes becomes large. We show that the required number of secure keys to be stored in the memory of each sensors is order of magnitude *smaller* than what is required for the Eschenauer-Gligor scheme [5].**

Keywords: Wireless sensor networks, Security, Key predistribution, Unassailability, Unsplittability.

I. Introduction

It is envisioned that security will constitute a key challenge for wireless sensor networks (WSNs) deployed in hostile environments. Unfortunately, many security schemes developed for general network environments do not take into account the unique features of WSNs: Public key cryptography is not computationally feasible because of the severe limitations imposed on the physical memory and power consumption of the individual sensors. Traditional key exchange and distribution protocols are based on trusting third parties, and this makes them inadequate for large-scale WSNs whose topologies are unknown prior to deployment. We refer the reader to [1], [5], [8] for discussions of the security challenges in WSN settings.

Random key predistribution schemes were introduced to address some of these difficulties. The idea of randomly assigning secure keys to sensor nodes prior to network deployment was first introduced by Eschenauer and Gligor [5]. Since then, many competing alternatives to the Eschenauer and Gligor (EG) scheme have been proposed; see [1] for a detailed survey of various key distribution schemes for WSNs. In this paper we consider the random pairwise key predistribution scheme of Chan et al. [2] and analyze its resiliency against sensor capture attacks. Interest in this scheme stems from the following advantages over the EG scheme: (i) Even if some nodes are captured, the secrecy of the remaining nodes is *perfectly* preserved; and (ii) Both node-to-node authentication and quorum-based revocation are enabled.

Given these advantages, we have found it of interest to model the pairwise scheme and to assess its performance. A number of issues related to secure connectivity and to the dimensioning of memory sizes have been discussed in the papers [9], [12]. In the present paper, we are interested in evaluating the resiliency of the pairwise scheme against node capture attacks.

The setup is as follows: An extremely powerful and knowledgeable adversary captures a number of sensors with the goal of severely impairing the functionality of the whole network. As was done in [7] for the EG scheme, the main question to be discussed here is whether this objective can be achieved by capturing a *small* number of sensors.

The analysis is given in the many node regime: We first look at the asymptotic behavior of the *maximum* number $C_r(n;K)$ of edges that can be compromised by capturing r nodes vs. the total number $|E(n;K)|$ of edges in the network as the number n of sensors grows unboundedly large – Here K is the parameter specifying the pairwise scheme; see Section II for details. Next, in the same regime we characterize the asymptotic behavior of the size $I_r(n;K)$ of the *largest* subset of sensors whose communications with the rest of the network can be compromised by capturing r nodes. For both quantities we give conditions on the scheme parameter and on r that ensure that if $r_n = o(n)$, then with *high probability* $C_{r_n}(n;K)$ (resp. $I_{r_n}(n;K)$) grows sub-linearly with $|E(n;K)|$ (resp. n). These conditions are highly desirable as they imply that an adversary cannot impair a considerable part of the network without capturing a considerable number of nodes. Both conditions were introduced in [7] under the names of *unassailability* and *unsplittability*, respectively, and used to evaluate the resiliency of the EG scheme; see Section III for details. As discussed in Sections IV and V, a comparison of our results with those of [7] shows that both properties can be achieved by the pairwise scheme with memory requirements which are order of magnitude *smaller* than that of the EG scheme. Proofs are available in Sections VI and VII.

A few words on notation and conventions in use: For sequences $a, b : \mathbb{N}_0 \rightarrow \mathbb{R}_+$, we write $a_n = o(b_n)$ as a shorthand for $\lim_{n\to\infty} \frac{a_n}{b_n} = 0$. On the other hand, $a_n = O(b_n)$ means that there exists $C > 0$ such that $a_n \leq C \cdot b_n$ for all n sufficiently large, whereas we write $a_n = \Omega(b_n)$ if there exists $c > 0$ such that $a_n \geq c \cdot b_n$ for all n sufficiently large. Throughout, we make use of the standard bounds

$$\binom{n}{r} \le \left(\frac{en}{r}\right)^r, \qquad \begin{array}{l} r = 1, \ldots, n \\ n = 1, 2, \ldots \end{array} \tag{1}$$

II. The model

The random pairwise key predistribution scheme of Chan et al. is parametrized by two positive integers n and K such that $K < n$. There are n nodes which are labeled $i = 1, \ldots, n$ with unique ids $\mathrm{Id}_1, \ldots, \mathrm{Id}_n$. Write $\mathcal{N} := \{1, \ldots n\}$ and set $\mathcal{N}_{-i} := \mathcal{N} - \{i\}$ for each $i = 1, \ldots, n$. With node i we associate a subset $\Gamma_{n,i}(K)$ of K nodes selected at *random* from $\mathcal{N}_{-i}$ – We say that each of the K nodes in $\Gamma_{n,i}(K)$ is paired to node i. Thus, for any subset $A \subseteq \mathcal{N}_{-i}$, we require

$$\mathbb{P}\left[\Gamma_{n,i}(K) = A\right] = \begin{cases} \binom{n-1}{K}^{-1} & \text{if } |A| = K \\ 0 & \text{otherwise} \end{cases},$$

ensuring that the selection of $\Gamma_{n,i}(K)$ is done *uniformly* amongst all subsets of $\mathcal{N}_{-i}$ which are of size K. Also, the set-valued rvs $\Gamma_{n,1}(K), \ldots, \Gamma_{n,n}(K)$ are assumed to be mutually independent.

Once this *offline* random pairing has been created, we construct the key rings $\Sigma_{n,1}(K), \ldots, \Sigma_{n,n}(K)$, one for each node, as follows: Assumed available is a collection of nK distinct cryptographic keys $\{\omega_{i|\ell},\ i = 1, \ldots, n;\ \ell = 1, \ldots, K\}$. Fix $i = 1, \ldots, n$ and let $\ell_{n,i} : \Gamma_{n,i}(K) \to \{1, \ldots, K\}$ denote a labeling of $\Gamma_{n,i}(K)$. For each node j in $\Gamma_{n,i}$ paired to i, the cryptographic key $\omega_{i|\ell_{n,i}(j)}$ is associated with j. For instance, if the random set $\Gamma_{n,i}(K)$ is realized as $\{j_1, \ldots, j_K\}$ with $1 \le j_1 < \ldots < j_K \le n$, then an obvious labeling consists in $\ell_{n,i}(j_k) = k$ with key $\omega_{i|k}$ associated with node j_k for each $k = 1, \ldots, K$. Finally, the pairwise key $\omega^\star_{n,ij} = [\mathrm{Id}_i | \mathrm{Id}_j | \omega_{i|\ell_{n,i}(j)}]$ is constructed and inserted in the memory modules of both nodes i and j. The key $\omega^\star_{n,ij}$ is assigned *exclusively* to the pair of nodes i and j, hence the terminology pairwise distribution scheme. The key ring of node i is the set

$$\Sigma_{n,i}(K) := \{\omega^\star_{n,ij}(K),\ j \in \Gamma_{n,i}(K)\} \cup \{\omega^\star_{n,ji},\ i \in \Gamma_{n,j}(K)\}.$$

If two nodes, say i and j, are within communication range of each other, they will be able to establish a secure edge if at least one of the events $i \in \Gamma_{n,j}$ or $j \in \Gamma_{n,j}$ is taking place – Both events may take place, in which case the memory modules of node i and j both contain the distinct keys $\omega^\star_{n,ij}$ and $\omega^\star_{n,ji}$.

Under full visibility, namely when every pair of nodes are within transmission range of each other, the pairwise scheme gives rise to the following class of random graphs: We say that the distinct nodes i and j are adjacent, written $i \sim j$, if and only if they have at least one key in common in their key rings, namely,

$$i \sim j \quad \text{iff} \quad \Sigma_{n,i}(K) \cap \Sigma_{n,j}(K) \neq \emptyset. \tag{2}$$

We denote by $\mathbb{H}(n;K)$ the undirected random graph on the vertex set $\{1, \ldots, n\}$ induced by the adjacency notion (2); this corresponds to modeling the pairwise distribution scheme under full visibility. Finally, let $E(n;K)$ denote the (random) set of edges in $\mathbb{H}(n;K)$.

III. Security metrics and resiliency

A. Measuring resiliency in WSNs

As we seek to understand the resiliency of the network against external attacks, we first specify the capabilities of the adversary considered here. To do so we adopt the following model already used in [7]: The adversary (sometimes also called the attacker), upon launching an attack against the network, captures some of its nodes, as a result of which it now owns the key rings stored at the captured nodes. An edge between two nodes is deemed *compromised* if the adversary owns a key which is stored in *both* their key rings. By the nature of the pairwise scheme this happens as soon as any one of the nodes has been captured. The adversary is assumed to have unlimited computing power; in particular it is expected to have sufficient knowledge of the network to minimize the number of nodes that need to be captured in order to compromise a given number of edges.

In many WSN applications, the network as a whole can still operate in a useful manner even though a *small* number of sensors have fallen under the control of the adversary [7]. In such situations it might be more relevant to protect the global functionality of the network rather than a few individual communication edges. However, if the adversary is capable of capturing a large fraction of the nodes, then there is not much that can be done to salvage the network functionalities. Hence, in evaluating the level of security provided by a key predistribution scheme, it is natural to ask whether *significant* damage to network functionalities can be inflicted by capturing just a small number of nodes. The next two sections provide ways to quantify this issue.

B. Unassailability

With A being the set of sensor nodes captured by the adversary, let $C_A(n;K)$ denote the total number of edges that are compromised as a result of this attack. In other words, $C_A(n;K)$ is the total number of edges (in the random graph $\mathbb{H}(n;K)$) with the property that at least one end of the edge is a node in A, i.e.,

$$C_A(n;K) = \left| \left\{ (i,j) : \begin{array}{c} 1 \le i < j \le n \\ i \sim j \end{array},\ i \in A \vee j \in A \right\} \right|.$$

The adversary under consideration is capable of maximizing $C_A(n;K)$ for a given number $|A|$ of nodes to be captured. This prompts us to introduce for each $r = 1, \ldots, n$, the maximum number $C_r(n;K)$ of edges that can be compromised by capturing r nodes, namely

$$C_r(n;K) := \max\left(C_A(n;K) :\ A \in \mathcal{N}_r\right)$$

where $\mathcal{N}_r$ denotes the collections of all subsets of $\{1, \ldots, n\}$ with exactly r elements.

Under the assumptions made on its capabilities, the powerful and knowledgeable attacker considered here will be able to compromise $C_r(n;K)$ edges by capturing (the appropriate) r nodes – This reflects a worst case mindset from the perspective of the network. Given this definition, it is natural to ask how the quantity $C_r(n;K)$ behaves in relation to the total number

$|E(n;K)|$ of edges as n gets large (with K and r also possibly scaled with n). It is common practice [3], [7] to regard the condition

$$C_{r_n}(n;K) = o(|E(n;K)|) \quad \text{whenever} \quad r_n = o(n) \tag{3}$$

as indicative of the resiliency of the network against node capture attacks. A crucial implication of the condition (3) is that in the many node regime, it implies that an adversary will not compromise $\Omega(|E(n;K_n)|)$ edges by taking over $o(n)$ nodes. We shall use condition (3) as a basis for characterizing the *unassailability* of the pairwise scheme. More specifically, we shall give conditions on K and n such that for any $\varepsilon > 0$, we have

$$\lim_{n\to\infty} \mathbb{P}\left[C_{r_n}(n;K) \geq \varepsilon \cdot |E(n;K)|\right] = 0 \tag{4}$$

whenever $r_n = o(n)$. When the parameter K is also scaled with n, the condition (4) will be used with K replaced by K_n.

C. Unsplittability

The metric (4) checks whether an adversary can compromise a considerable fraction of edges by launching an attack on few sensors. But, it does not tell anything about the ability of the adversary to *disconnect* the network. To explore this issue further, with A still acting as the set of nodes taken over by the attacker, we say that the subset S of nodes is A-*splittable* if the adversary can compromise all the edges from S to $S^c = \mathcal{N} - S$ by capturing the nodes in A. To be more precise, for any subset S of nodes let $E(n;K)(S)$ denote the set of edges in $\mathbb{H}(n;K)$ with one end in S and the other in S^c. Then, the A-splittability of S is characterized by

$$\wedge_{(i,j)\in E(n;K)(S)} \left(i \in A \vee j \in A\right). \tag{5}$$

This is because once the set of nodes in A captured, an edge $i \sim j$ in $\mathbb{H}(n;K)$ will be compromised if either condition $i \in A$ or $j \in A$ takes place.

Given the infinite computational power available to it, the attacker can in principle minimize the number of nodes it needs to capture in order to *split* S from the rest of the network. Thus, for each $r = 1, \ldots, n-1$, we say that the set S of nodes is r-*splittable* whenever there *exists* a set A of r nodes such that S is A-splittable. The r-splittability of S is encoded through the conditions

$$\vee_{A\in\mathcal{N}_r} \left(\wedge_{(i,j)\in E(n;K)(S)} \left(i \in A \vee j \in A\right)\right). \tag{6}$$

It is clear that if S is r-splittable, then its complement S^c (in $\mathcal{N}$) is also r-splittable. Finally, let $I_r(n;K)$ denote the size of the largest subset S (with size $|S| \leq \frac{n}{2}$) that can be disconnected from the rest of the network by capturing r nodes, namely

$$I_r(n;K) = \max\left\{|S| : S \subseteq \mathcal{N}, |S| \leq \frac{n}{2},\ S \text{ is } r\text{-splittable}\right\}.$$

It is natural to wonder as to the behavior of $I_r(n;K)$ as n grows large – It is always the case that $r \leq I_r(n;K) \leq \frac{n}{2}$. From the perspective of the network, it is desirable that the largest subset which can be disconnected be small whenever the number of captured nodes is small. As in [7] this leads to the condition

$$I_{r_n}(n;K) = o(n) \quad \text{whenever} \quad r_n = o(n)$$

as our second characterization of resiliency. In this paper, we give conditions on how to scale K with the number n of nodes such that for any $0 < \gamma \leq \frac{1}{2}$, we have

$$\lim_{n\to\infty} \mathbb{P}\left[I_{r_n}(n;K_n) \geq \gamma n\right] = 0 \tag{7}$$

whenever $r_n = o(n)$ – From these definitions it follows that (7) holds trivially when $\gamma > \frac{1}{2}$. The operational usefulness of (7) derives from the fact that it ensures that for any subset S of $\mathcal{N}$, with $|S| = \Omega(n)$, an adversary must capture *at least* $\Omega(n)$ nodes in order to compromise *all* edges from S to S^c.

IV. Relevant prior work

The resiliency of WSNs against node capture attacks was also investigated by Mei et al. [7]: They considered the EG scheme as the underlying security mechanism and obtained conditions on the scheme parameters to ensure the appropriate analogs of (4) and (7). We now summarize their findings in order to identify the number of keys (to be kept in the memory of each sensor) that is required to ensure the desired conditions (4) and (7).

Let $\mathbb{K}(n;\theta)$ denote the random key graph on the vertex set $\{1,\ldots,n\}$ induced by the EG scheme under full visibility [13]; here $\theta = (\Sigma_{\rm EG}, P)$ collectively stands for the parameters that specify the EG scheme, namely the (fixed) size $\Sigma_{\rm EG}$ of the key ring of each sensor node and the size P of the key pool. Thus, let $\Sigma_{n,1}(\theta), \ldots, \Sigma_{n,n}(\theta)$ denote the key rings associated with nodes $1,\ldots,n$, respectively, in the EG scheme. By construction, $|\Sigma_{n,1}(\theta)| = \cdots = |\Sigma_{n,n}(\theta)| := \Sigma_{\rm EG}$. We are now in a position to present the main result obtained in [7]. A scaling for the EG scheme is any pair of mappings $\Sigma_{\rm EG}, P : \mathbb{N}_0 \to \mathbb{N}_0$ such that

$$\Sigma_{{\rm EG},n} \leq P_n, \quad n = 2, 3, \ldots$$

Theorem 4.1: Consider any scaling $\Sigma_{\rm EG}, P : \mathbb{N}_0 \to \mathbb{N}_0$ *for the EG scheme which satisfies*

$$\Sigma_{{\rm EG},n} \geq \sqrt{n \log n}. \tag{8}$$

Then, (4) and (7) hold.

In [7] it is claimed, but without proofs, that both properties hold also when $\Sigma_{{\rm EG},n} \geq \log n$. The condition (8) was derived so as to also ensure that $\mathbb{K}(n;\theta_n)$ is asymptotically almost surely (a.a.s) connected. Here, to comply with that practice, we recall sufficient conditions for $\mathbb{H}(n;K)$ to be a.a.s. connected. To fix the terminology, we refer to any mapping $K : \mathbb{N}_0 \to \mathbb{N}_0$ as a *scaling* (for the pairwise scheme) provided

$$K_n < n, \quad n = 2, 3, \ldots$$

In [11], the following was shown:

Theorem 4.2: For any scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$ *such that* $K_n \geq 2$ *for all* n *sufficiently large, it holds that* $\lim_{n\to\infty} \mathbb{P}\left[\mathbb{H}(n;K_n) \text{ is connected}\right] = 1$.

V. Main results and discussion

The main result of the paper, given next, provides a version of Theorem 4.1 for the pairwise scheme.

Theorem 5.1: Consider any scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$. We always have (4), whereas (7) is satisfied whenever

$$\lim_{n\to\infty} K_n = \infty. \tag{9}$$

Theorem 5.1, which is established in Section VII, gives conditions for unassailability and unsplittability under the pairwise scheme. However, in contrast with the EG scheme and its variants, the key rings $\Sigma_{n,1}(K), \ldots, \Sigma_{n,n}(K)$ produced by the pairwise scheme are of variable size between K and $K+(n-1)$. Therefore, in order to meaningfully compare our findings with those for the EG scheme from [7], we need to understand how the sizes $|\Sigma_{n,1}(K)|, \ldots, |\Sigma_{n,n}(K)|$ of these key rings depend on K and n.

To explore this issue further, observe that

$$|\Sigma_{n,i}(K)| = K + \sum_{j=1,\ j\neq i}^{n} \mathbf{1}\left[i \in \Gamma_{n,j}(K)\right], \quad i = 1, \ldots, n$$

so that

$$|\Sigma_{n,i}(K)| =_{st} K + \text{Bin}\left(n-1, K/(n-1)\right), \tag{10}$$

whence $\mathbb{E}\left[|\Sigma_{n,i}(K)|\right] = 2K$. Since every key appears in exactly two different key rings it follows that

$$|\Sigma|_{n,\text{Avg}}(K) := \frac{|\Sigma_{n,1}(K)| + \cdots + |\Sigma_{n,n}(K)|}{n} = 2K$$

by construction. Furthermore, in order to deal with worst case scenarios, we introduce the maximal key ring size given by

$$|\Sigma|_{n,\text{Max}}(K) := \left(\max_{i=1,\ldots,n} |\Sigma_{n,i}(K)|\right), \quad n = 2, 3, \ldots.$$

Next, upon using a standard Hoeffding bound [4, Thm. 1.1, p. 6] for the binomial rvs (10), we obtain the following concentration result for the maximal key ring size. This result can be established with the help of standard bounding arguments, but is omitted here due to space limitations.

Theorem 5.2: Consider any scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$ such that $K_n = O(\log n)$. Then, there exists $c > 0$ such that

$$\lim_{n\to\infty} \mathbb{P}\left[|\Sigma|_{n,\text{Max}}(K_n) > cK_n\right] = 0. \tag{11}$$

In view of Theorem 4.1 and Theorem 5.1, we can now compare the security properties of the pairwise scheme and of the EG scheme. It is clear from Theorem 5.1 and (11) that the pairwise key distribution scheme can ensure (4) with *all* key rings being on the order $\log n$. Similarly, Theorem 5.1 and (11) show that to ensure unsplittability, the pairwise scheme requires key ring sizes of $O(\log n)$. As we compare these findings with Theorem 4.1, we see that the pairwise scheme can achieve both properties with much smaller key ring sizes than needed for the EG scheme; see Figure 1.

	Unassailability	Unsplittability
EG – Σ_{EG}	$\Omega(\sqrt{n \log n})$	$\Omega(\sqrt{n \log n})$
Pairwise – $\lvert\Sigma\rvert_{\text{Avg}}$	4	w_n
Pairwise – $\lvert\Sigma\rvert_{\text{Max}}$	$O(\log n)$	$O(\log n)$

Fig. 1. *A comparison of the EG scheme and the pairwise scheme in terms of the minimum number of keys required to achieve unassailability and unsplittability. As before, w_n stands for any function satisfying $\lim_{n\to\infty} w_n = \infty$. It is clear that pairwise scheme can ensure both of the desired properties with much less memory load on the sensors as compared to the EG scheme.*

VI. A basic inequality

Both assertions in Theorem 5.1 are established in Section VII, and rely on a basic inequality discussed next. For every $\varepsilon > 0$ and $K = 1, 2, \ldots$, set

$$H_\varepsilon(x; K) = (\varepsilon - x)K \log 2 + x \log\left(\frac{x}{e}\right), \quad 0 \le x \le 1$$

Proposition 6.1: With $\varepsilon > 0$, consider positive integers K and n such that $K < n$. Then, for each $r = 1, 2, \ldots, n$, we have

$$\mathbb{P}\left[C_r(n; K) \ge \varepsilon n K\right] \le e^{-nH_\varepsilon\left(\frac{r}{n}; K\right)} \tag{12}$$

whenever

$$\varepsilon > \frac{r}{n}\left(1 + 2e\frac{n-r}{n-1}\right). \tag{13}$$

Proof. Pick a subset A of nodes. The exact expression

$$\begin{aligned} C_A(n; K) \quad = \quad & \frac{1}{2}\sum_{i\in A}\sum_{j\in A} \mathbf{1}\left[j \in \Gamma_{n,i}(K) \vee i \in \Gamma_{n,j}(K)\right] \\ & + \sum_{i\in A}\sum_{k\in A^c} \mathbf{1}\left[k \in \Gamma_{n,i}(K) \vee i \in \Gamma_{n,k}(K)\right] \end{aligned}$$

is easily established but cumbersome to work with. Instead we will rely on the bound

$$C_A(n; K) \le |A|K + E_{n,A}(K) \tag{14}$$

where we have set

$$E_{n,A}(K) := \sum_{j\in A^c}\sum_{i\in A} \mathbf{1}\left[i \in \Gamma_{n,j}(K)\right].$$

The validity of (14) can be seen as follows: There are at most $K|A|$ compromised edges originating out of nodes in A, while there are exactly $E_{n,A}(K)$ compromised edges originating out of nodes in A^c. To simplify the notation we shall write $E_{n,A}(K) = E_{n,r}(K)$ when $A = \{1, \ldots, r\}$ with $r = 1, \ldots, n$.

Now fix $r = 1, \ldots, n$ and $\varepsilon > 0$. Using (14) we find

$$\begin{aligned}
&\mathbb{P}\left[C_r(n;K) \geq \varepsilon nK\right] \\
&\quad = \quad \mathbb{P}\left[\bigcup_{A \in \mathcal{N}_r} \left[C_A(n;K) \geq \varepsilon nK\right]\right] \\
&\quad \leq \quad \mathbb{P}\left[\bigcup_{A \in \mathcal{N}_r} \left[E_{n,A}(K) \geq \varepsilon nK - rK\right]\right] \\
&\quad \leq \quad \sum_{A \in \mathcal{N}_r} \mathbb{P}\left[E_{n,A}(K) \geq \varepsilon nK - rK\right] \\
&\quad = \quad \binom{n}{r} \mathbb{P}\left[E_{n,r}(K) \geq \varepsilon nK - rK\right]. \qquad (15)
\end{aligned}$$

In [10] the rvs $\{\mathbf{1}\left[i \in \Gamma_{n,j}(K)\right],\ j = r+1, \ldots, n;\ i = 1, \ldots, r\}$ were shown to be negatively associated [6]. As a result, the Chernoff-Hoeffding bound [4, Thm. 1.1, p. 6] applies to the sum $E_{n,r}(K)$ in the form

$$\mathbb{P}\left[E_{n,r}(K) \geq t\right] \leq 2^{-t} \qquad (16)$$

whenever $t > 0$ satisfies

$$t > 2e \cdot \mathbb{E}\left[E_{n,r}(K)\right] = 2e \cdot r(n-r)\frac{K}{n-1} \qquad (17)$$

since $\mathbb{E}\left[E_{n,r}(K)\right] = r(n-r)\frac{K}{n-1}$. Note that (17) with $t = \varepsilon nK - rK$ is equivalent to (13), in which case (16) becomes

$$\mathbb{P}\left[E_{n,r}(K) \geq \varepsilon nK - rK\right] \leq e^{-K(\varepsilon n - r)\log 2}. \qquad (18)$$

Reporting the bounds (1) and (18) into (15), the conclusion (12) readily follows. ■

VII. A proof of Theorem 5.1

Consider the random graph $\mathbb{H}(n;K)$ for positive integers n and K such that $K < n$. By construction each key is associated with one and only one edge in $\mathbb{H}(n;K)$, whereas at most two keys can be associated with a given edge. Thus, for edge $i \sim j$, the upper bound is reached when both events $i \in \Gamma_{n,i}(K)$ and $j \in \Gamma_{n,i}(K)$ take place. As a result, we have

$$\frac{Kn}{2} \leq |E(n;K)| \leq Kn. \qquad (19)$$

Now consider any scaling $K : \mathbb{N}_0 \rightarrow \mathbb{N}_0$ and assume that the condition

$$r_n = o(n) \qquad (20)$$

holds. Given $\varepsilon > 0$, the condition

$$\frac{\varepsilon}{2} > \frac{r_n}{n}\left(1 + 2e\frac{n - r_n}{n-1}\right) \qquad (21)$$

will be met for all n sufficiently large. On that range, Proposition 6.1 (with ε replaced by $\frac{\varepsilon}{2}$) yields

$$\mathbb{P}\left[C_{r_n}(n;K_n) \geq \varepsilon\frac{nK_n}{2}\right] \leq e^{-nH_{\frac{\varepsilon}{2}}(\frac{r_n}{n};K_n)} \qquad (22)$$

and the convergence

$$\lim_{n\to\infty} \mathbb{P}\left[C_{r_n}(n;K_n) \geq \varepsilon\frac{nK_n}{2}\right] = 0 \qquad (23)$$

follows since $\liminf_{n\to\infty} H_{\frac{\varepsilon}{2}}\left(\frac{r_n}{n};K_n\right) > 0$ under (20). The desired conclusion (4) is obtained from (23) upon using (19).■

As we now turn to establishing (7), fix the positive integers n and K such that $K < n$. The discussion starts with the following observation: Consider an attack that succeeds in capturing the nodes in A, and let S denote an arbitrary subset of nodes. If S is A-splittable, then all the edges between the set of nodes S and its complement S^c are compromised by the capture of nodes in A. Hence, the total number $C_A(n;K)$ of edges which are compromised by this attack must be at least $|E(n;K)(S)|$. Therefore, by the characterization (5) of S being A-splittable we have the inclusion

$$[S \text{ is } A\text{-splittable}] \subseteq [C_A(n;K) \geq |E(n;K)(S)|].$$

For each γ in $\left(0, \frac{1}{2}\right]$, let $\mathcal{N}_{n,\gamma}$ denote the collection of all subsets S of $\mathcal{N}$ such that $\gamma n \leq |S| \leq \frac{n}{2}$. For each $r = 1, \ldots, n$, the definition of the count variable $I_r(n;K)$ and this last inclusion imply

$$\begin{aligned}
&\mathbb{P}\left[I_r(n;K) \geq \gamma n\right] \\
&\quad = \quad \mathbb{P}\left[\bigcup_{S \in \mathcal{N}_{n,\gamma}} [S \text{ is } r\text{-splittable}]\right] \\
&\quad = \quad \mathbb{P}\left[\bigcup_{S \in \mathcal{N}_{n,\gamma}} \bigcup_{A \in \mathcal{N}_r} [S \text{ is } A\text{-splittable}]\right] \\
&\quad \leq \quad \mathbb{P}\left[\bigcup_{S \in \mathcal{N}_{n,\gamma}} \bigcup_{A \in \mathcal{N}_r} [C_A(n;K) \geq |E(n;K)(S)|]\right] \\
&\quad = \quad \mathbb{P}\left[\bigcup_{S \in \mathcal{N}_{n,\gamma}} [C_r(n;K) \geq |E(n;K)(S)|]\right] \\
&\quad \leq \quad \sum_{S \in \mathcal{N}_{n,\gamma}} \mathbb{P}\left[C_r(n;K) \geq |E(n;K)(S)|\right] \qquad (24)
\end{aligned}$$

upon using a union bound argument in the last step.

Next, pick $\varepsilon > 0$ and δ in $(0,1)$ such that

$$2\varepsilon < (1-\delta)\gamma. \qquad (25)$$

The need for doing so will become apparent below. For each S in $\mathcal{N}_{n,\gamma}$, conditioning on $|E(n;K)(S)| \geq \varepsilon nK$ yields

$$\begin{aligned}
&\mathbb{P}\left[C_r(n;K) \geq |E(n;K)(S)|\right] \qquad (26) \\
&\quad \leq \quad \mathbb{P}\left[C_r(n;K) \geq \varepsilon nK\right] + \mathbb{P}\left[|E(n;K)(S)| < \varepsilon nK\right].
\end{aligned}$$

If condition (13) *were* to hold, then Proposition 6.1 would imply

$$\begin{aligned}
&\sum_{S \in \mathcal{N}_{n,\gamma}} \mathbb{P}\left[C_r(n;K) \geq \varepsilon nK\right] \\
&\quad \leq \quad |\mathcal{N}_{n,\gamma}| \cdot e^{-nH_\varepsilon(\frac{r}{n};K)} \\
&\quad \leq \quad e^{-n\left(H_\varepsilon(\frac{r}{n};K) - \log 2\right)} \qquad (27)
\end{aligned}$$

since $|\mathcal{N}_{n,\gamma}| \leq 2^n$.

As we consider the second term in the right handside of (26), pick S in $\mathcal{N}_{n,\gamma}$ and observe that

$$\begin{aligned} |E(n;K)(S)| &= \sum_{j\in S^c}\sum_{i\in S} \mathbf{1}\left[j \in \Gamma_{n,i}(K) \vee i \in \Gamma_{n,j}(K)\right] \\ &\geq E_{n,S}(K). \end{aligned} \tag{28}$$

As before, the negative association of the rvs $\{\mathbf{1}\left[i \in \Gamma_{n,j}(K)\right], \ i \in S, j \in S^c\}$, shown in [10], validates the Chernoff-Hoeffding bound for the sum $E_{n,S}(K)$ [4, Thm. 1.1, p. 6] in the form

$$\mathbb{P}\left[E_{n,S}(K) \leq (1-\delta)\mathbb{E}\left[E_{n,S}(K)\right]\right] \leq e^{-\frac{\delta^2}{2}\mathbb{E}[E_{n,S}(K)]}. \tag{29}$$

Note also that

$$\mathbb{E}\left[E_{n,S}(K)\right] = |S|\left(n-|S|\right)\cdot\frac{K}{n-1} \geq \frac{\gamma}{2}\cdot nK$$

since $\gamma n \leq |S| \leq \frac{n}{2}$ by membership of S in $\mathcal{N}_{n,\gamma}$. From (25) we automatically have

$$\varepsilon nK < (1-\delta)\mathbb{E}\left[E_{n,S}(K)\right] \tag{30}$$

for *all* $n = 1, 2, \ldots$. Using the bounds (28) and (30) together with (29), we conclude

$$\begin{aligned} &\sum_{S\in\mathcal{N}_{n,\gamma}} \mathbb{P}\left[|E(n;K)(S)| < \varepsilon nK\right] \\ &\leq \sum_{S\in\mathcal{N}_{n,\gamma}} \mathbb{P}\left[E_{n,S}(K) < \varepsilon nK\right] \\ &\leq \sum_{S\in\mathcal{N}_{n,\gamma}} \mathbb{P}\left[E_{n,S}(K) < (1-\delta)\mathbb{E}\left[E_{n,S}(K)\right]\right] \\ &\leq \sum_{S\in\mathcal{N}_{n,\gamma}} e^{-\frac{\delta^2}{2}\mathbb{E}[E_{n,S}(K)]} \\ &\leq \sum_{S\in\mathcal{N}_{n,\gamma}} e^{-\frac{\delta^2}{2}\cdot\frac{\gamma}{2}nK} \\ &\leq \left(2e^{-\gamma\frac{\delta^2}{4}\cdot K}\right)^n. \end{aligned} \tag{31}$$

Consider now a scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$ satisfying (9) and replace K by K_n for all $n = 1, 2, \ldots$, possibly making r depend on n as well. As in the earlier part of the proof, under (20) the condition (21) (with r replaced by r_n) holds for all $n = 1, 2, \ldots$ sufficiently large, whence (27) holds on that range. It is now plain that

$$\lim_{n\to\infty} \sum_{S\in\mathcal{N}_{n,\gamma}} \mathbb{P}\left[C_{r_n}(n;K_n) \geq \varepsilon nK_n\right] = 0$$

since $\lim_{n\to\infty}\left(H_\varepsilon(\frac{r_n}{n};K_n) - \log 2\right) = \infty$ under the conditions (9) and (20). Similarly it is plain from (31) that

$$\lim_{n\to\infty} \sum_{S\in\mathcal{N}_{n,\gamma}} \mathbb{P}\left[|E(n;K_n)(S)| < \varepsilon nK_n\right] = 0.$$

The desired conclusion (7) is now an easy consequence of the last two convergence statements when coupled with the bounds (24) and (26). ■

ACKNOWLEDGMENT

This work was supported by NSF Grant CCF-07290.

REFERENCES

[1] S. A. Çamtepe and B. Yener, *Key Distribution Mechanisms for Wireless Sensor Networks: a Survey,* Technical Report TR-05-07, Rensselaer Polytechnic Institute, Computer Science Department, Troy (NY), March 2005.

[2] H. Chan, A. Perrig, D. Song, “Random key predistribution schemes for sensor networks,” Proceedings of SP 2003, Oakland (CA), May 2003.

[3] W. Du, J. Deng, Y.S. Han and P.K. Varshney, “A pairwise key predistribution scheme for wireless sensor networks,” Proceedings of CCS 2003, October 2003.

[4] D. Dubhashi and A. Panconesi, *Concentration of Measure for the Analysis of Randomized Algorithms*, Cambridge University Press, New York (NY), 2009.

[5] L. Eschenauer and V.D. Gligor, “A key-management scheme for distributed sensor networks,” Proceedings of CCS 2002, pp. 41-47.

[6] K. Joag-Dev and F. Proschan, “Negative association of random variables, with applications,” *The Annals of Statistics* **11** (1983), pp. 266-295

[7] A. Mei, A. Panconesi and J. Radhakrishnan, “Unassailable sensor networks,” Proceedings of SecureComm 2008, September 2008.

[8] A. Perrig, J. Stankovic and D. Wagner, “Security in wireless sensor networks,” *Communications of the ACM* **47** (2004), pp. 53–57.

[9] O. Yağan and A. M. Makowski, “On the gradual deployment of random pairwise key distribution schemes,” Proceedings of WiOpt 2011, Princeton (NJ), May 2011.

[10] O. Yağan and A. M. Makowski, “Modeling the pairwise key distribution scheme in the presence of unreliable links.” Available online at `arXiv:1102.2250v1[cs.IT]`.

[11] O. Yağan and A. M. Makowski, “On random pairwise graphs.” Available online at `http://www.ece.umd.edu/˜oyagan/Journals/Pairwise-DM.pdf`

[12] O. Yağan and A. M. Makowski, “Designing securely connected wireless sensor networks in the presence of unreliable links,” in Proceedings of the IEEE International Conference on Communications (ICC 2011), Kyoto (Japan), June 2011.

[13] O. Yağan, *Random Graph Modeling of Key Distribution Schemes in Wireless Sensor Networks*, Ph.D. Thesis, Department of Electrical and Computer Engineering, University of Maryland, College Park (MD), June 2011.