Contents lists available at SciVerse ScienceDirect

Performance Evaluation

journal homepage: www.elsevier.com/locate/peva

ELSEVIER

On the scalability of the random pairwise key predistribution scheme: Gradual deployment and key ring sizes

Osman Yağan, Armand M. Makowski *

Department of Electrical and Computer Engineering and the Institute for Systems Research, University of Maryland, College Park, MD 20742, USA

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Wireless sensor networks
Security
Key predistribution
Random key graphs
Connectivity
Memory requirements

ABSTRACT

The pairwise key distribution scheme of Chan et al. is a randomized key predistribution scheme which enables cryptographic protection in wireless sensor networks (WSNs). Although this pairwise scheme has several advantages over other randomized key predistribution schemes, including the original scheme of Eschenauer and Gligor, it has been deemed *non*-scalable for the following two reasons: (i) There are implementation difficulties when sensors need to be deployed in multiple stages; and (ii) the possibly large number of keys stored at each sensor node is random and may vary from sensor to sensor. Here, we explore these issues as follows: (i) We propose an implementation of the pairwise scheme that supports the gradual deployment of sensor nodes in several consecutive phases; (ii) We show how the scheme parameters should scale with the maximum number n of sensor nodes in the network so that a.a.s. secure connectivity is maintained throughout every deployment phase of the network; (iii) We derive scaling conditions on the scheme parameters for *all* the sensors in the network to have $O(\log n)$ many keys with very high probability as n grows large. This dimensioning of the key rings can be made to support a.a.s. secure connectivity at every step of deployment.

1. Introduction

1.1. Security in WSNs

Wireless sensor networks (WSNs) are distributed collections of sensors with limited capabilities for computations and wireless communications. Such networks are likely to be deployed in hostile environments where cryptographic protection will be needed. Providing this cryptographic protection has been identified as a serious challenge to the successful deployment of WSNs. However, traditional key exchange and distribution protocols based on trusting third parties have been found inadequate for large-scale WSNs; see the references [1–4] for discussions of some of the obstacles.

Random key predistribution schemes were recently introduced to address some of the difficulties. The idea of randomly assigning secure keys to sensor nodes prior to network deployment was first proposed by Eschenauer and Gligor [1]. The EG scheme, as we refer to it hereafter, has been extensively investigated [1,5–11], with most of the focus being on the *full visibility* case where nodes are all within communication range of each other. Although the assumption of full visibility does away with the wireless nature of the communication medium supporting WSNs, this simplification does make it possible to understand how randomizing the key selections affects the establishment of a secure network in the best of circumstances.

The work of Eschenauer and Gligor has spurred the development of other key distribution schemes which perform better than the EG scheme in some respects, e.g., see [2,3,12,13]. Here we consider the random *pairwise* key predistribution scheme

[*] Corresponding author. Tel.: +1 301 405 6844; fax: +1 301 314 9281.
E-mail addresses: osmanyagan@gmail.com (O. Yağan), armand@isr.umd.edu (A.M. Makowski).

0166-5316/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.peva.2013.03.001

proposed by Chan et al. [12]: Before deployment, each of the n sensor nodes is paired (offline) with K distinct nodes which are randomly selected amongst all other $n-1$ nodes. For each such pairing, a *unique* pairwise key is generated and stored in the memory modules of each of the paired sensors along with the id of the other node—All keys are assumed to be of the same length. A secure link can then be established between two communicating nodes if at least one of them has been paired to the other (in which case they have at least one key in common). See Section 2 for implementation details. This scheme has the following advantages over the EG scheme (and others): (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.

1.2. Earlier work on random pairwise key distribution schemes

Given these advantages, we have found it of interest to assess the performance of the pairwise scheme, and have begun a formal investigation along these lines. Highlights of this earlier work include:

Connectivity under full visibility. Let $\mathbb{H}(n; K)$ denote the random graph on the vertex set $\{1, \ldots, n\}$ where distinct nodes i and j are adjacent if they have a pairwise key in common; as in earlier work on the EG scheme this corresponds to modeling the random pairwise distribution scheme under full visibility. In [14,15] we showed that the probability of $\mathbb{H}(n; K)$ being connected approaches 1 (resp. 0) as n grows large if $K \geq 2$ (resp. if $K = 1$), i.e., $\mathbb{H}(n; K)$ is asymptotically almost surely (a.a.s.) connected whenever $K \geq 2$. See Theorem 2.1 for details. □

Connectivity under partial visibility. In [16,17], the connectivity issue of the pairwise distribution scheme was revisited under more realistic assumptions that account for the possibility that communication links between nodes may not be available. This was done in the context of a simple communication model where the communication channels are mutually independent, and are either on or off. The appropriate random graph structure can be viewed as the intersection of the random graph $\mathbb{H}(n; K)$ with an Erdős–Rényi graph. In that framework we have established zero–one laws for two basic (and related) graph properties, namely graph connectivity and the absence of isolated nodes, when the model parameters are scaled with the number of users. The critical thresholds for both properties were identified and shown to coincide. □

1.3. Contributions

In spite of the aforementioned advantages over the original EG scheme and other randomized key predistribution schemes, the pairwise key distribution scheme of Chan et al. has been deemed *non*-scalable for the following two main reasons: (i) There are implementation difficulties when sensors need to be deployed in multiple stages; and (ii) The possibly large number of keys at a sensor node is random and may vary from sensor to sensor. These issues are addressed here to some extent, with the various contributions summarized below. A preliminary version of this work was presented with partial proofs in the conference paper [18] under a different title.

Connectivity under gradual deployment. In the present paper, we continue our study of connectivity properties for the scheme of Chan et al., still under full visibility, but from a different perspective: In many settings, e.g., environmental monitoring and military theaters, sensor nodes are expected to be deployed *gradually* over time in response to the demands of a dynamically changing situation: Data collected in an initial phase might suggest monitoring other regions, and this calls for the deployment of additional sensors. Yet, the pairwise key distribution is an *offline* pairing mechanism which simultaneously involves all n nodes. Thus, once the maximal number n of nodes to be deployed is set, there is no way to add more nodes to the network and still *recursively* expand the pairwise distribution scheme (as is possible for the EG scheme). However, as explained in Section 2.2, the gradual deployment of a large number of sensor nodes is nevertheless feasible from a practical viewpoint. In that context we are interested in understanding how the parameter K needs to scale with n large in order to ensure that connectivity is *maintained* a.a.s. throughout gradual deployment.

The main contributions along these lines take the following form: With $0 < \gamma < 1$, let $\mathbb{H}_\gamma(n; K)$ denote the subgraph of $\mathbb{H}(n; K)$ restricted to the nodes $1, \ldots, \lfloor \gamma n \rfloor$. We first present scaling laws for the absence of isolated nodes in $\mathbb{H}_\gamma(n; K)$ in the form of a full zero–one law, and use these results to formulate conditions under which $\mathbb{H}_\gamma(n; K)$ is a.a.s. *not* connected. Next, we consider a gradual deployment scenario with ℓ consecutive phases: With $0 < \gamma_1 < \gamma_2 < \cdots < \gamma_\ell < 1$, we assume that $\lfloor \gamma_k n \rfloor - \lfloor \gamma_{k-1} n \rfloor$ new nodes are deployed in the kth phase with $k = 1, \ldots, \ell$. We give conditions on n, K and γ_1 so that $\mathbb{H}_{\gamma_k}(n; K)$ is a.a.s. connected for each $k = 1, 2, \ldots, \ell$. This corresponds to the network being connected in *each* of the ℓ phases of the gradual deployment. □

Needed key ring sizes. The key rings for the pairwise scheme have variable size between K and $K + (n-1)$, in sharp contrast with the EG scheme (and its variants) where the key rings have all the same fixed size. This naturally brings up the question as to how many keys need to be kept in the memory module of a sensor in order to achieve certain desired properties, e.g., secure connectivity. Since sensor nodes are expected to have very limited memory, it is crucial for a key distribution scheme to have *low* memory requirements in order to be competitive [13,19]. With this in mind we first give minimal conditions on a scaling K_n such that the size of any key ring hovers around its mean $2K_n$ (in some precise probabilistic sense as n becomes large). Next, we sharpen this result by showing that the *maximum* key ring size is on the order $O(\log n)$ with very high probability provided $K_n = O(\log n)$. It turns out that taking $K_n = O(\log n)$ is sufficient to achieve secure connectivity under the pairwise scheme when the network is deployed in a single phase; see Theorem 2.1 for a much stronger result. In the gradual deployment scenario, we also show that the sensor network can maintain a.a.s. connectivity throughout all the phases of its deployment with $O(\log n)$ keys stored in each sensor's memory. Such a key ring size can be realized with very high probability by selecting $K_n = \Theta(\log n)$, and is comparable to that of the EG scheme (in realistic WSN scenarios [6]). This points to the possibility of turning the pairwise scheme into a competitive random key distribution scheme. □

As with the results in [15], the assumption of full visibility may yield a dimensioning of the pairwise scheme which is too optimistic. This is due to the fact that the unreliable nature of wireless links has not been incorporated in the model. However, the results given in this paper already yield a number of interesting observations: The zero–one laws obtained here differ significantly from the corresponding results in the single deployment case [15], and are not implied by them. Thus, the gradual deployment may have a significant impact on the dimensioning of the pairwise distribution algorithm. Yet, the required number of keys to achieve secure connectivity being $O(\log n)$, it is still feasible to use the pairwise scheme under gradual deployment; note that, the required key ring size in EG scheme is also $O(\log n)$ under full visibility [6].

1.4. Miscellaneous

The rest of the paper is organized as follows: In Section 2 we introduce a framework to model the random pairwise distribution scheme; this section also describes an implementation of the scheme which supports gradual network deployment. Section 3 presents results concerning connectivity and the absence of isolated nodes under gradual deployment; the proofs are available in Sections 6 and 7. Results on key ring sizes are stated in Section 4 and established in Section 8. Limited simulation experiments are discussed in Section 5.

A word on the notation used throughout the paper: All limiting statements, including asymptotic equivalences, are understood with n going to infinity. The random variables (rvs) under consideration are all defined on the same probability triple $(\Omega, \mathcal{F}, \mathbb{P})$. Probabilistic statements are made with respect to this probability measure $\mathbb{P}$, and we denote the corresponding expectation operator by $\mathbb{E}$. Also, we use the notation $=_{st}$ to indicate distributional equality. The indicator function of an event E is denoted by $\mathbf{1}[E]$. For any discrete set S we write $|S|$ for its cardinality.

2. The model

2.1. Implementing pairwise key distribution schemes

The random pairwise key predistribution scheme 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 distinct nodes selected uniformly and at random from $\mathcal{N}_{-i}$—Each of the K nodes in $\Gamma_{n,i}(K)$ is said to be *paired* to node i. Thus, for any subset $A \subseteq \mathcal{N}_{-i}$, we require

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

In other words, the selection of $\Gamma_{n,i}(K)$ is done *uniformly* amongst all subsets of $\mathcal{N}_{-i}$ which are of size exactly K. The rvs $\Gamma_{n,1}(K), \ldots, \Gamma_{n,n}(K)$ are assumed to be mutually independent so that

$$\mathbb{P}\left[\, \Gamma_{n,i}(K) = A_i,\ i = 1, \ldots, n \,\right] = \prod_{i=1}^{n} \mathbb{P}\left[\, \Gamma_{n,i}(K) = A_i \,\right]$$

for arbitrary $A_1, \ldots, A_n$ subsets of $\mathcal{N}_{-1}, \ldots, \mathcal{N}_{-n}$, respectively.

On the basis of this *offline* random pairing, we now 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 $\{\omega_{i|\ell},\ i = 1, \ldots, n;\ \ell = 1, \ldots, K\}$—These keys are drawn from a very large pool of keys; in practice the pool size is assumed to be much larger than nK, and can be safely taken to be infinite for the purpose of our discussion. It also assumed throughout that the keys used here are all of the *same* length.

Now, 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 $\Gamma_{n,i}(K)$ 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 $1 \le j_1 < \cdots < j_K \le n$, then an obvious labeling consists of $\ell_{n,i}(j_k) = k$ for each $k = 1, \ldots, K$ with key $\omega_{i|k}$ associated with node j_k. Of course other labelings are possible, e.g., according to decreasing labels or according to a random permutation. 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. Inherent to this construction is the fact that 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 $\Sigma_{n,i}(K)$ of node i is the set

$$\Sigma_{n,i}(K) := \left\{\omega^\star_{n,ij},\ j \in \Gamma_{n,i}(K)\right\} \cup \left\{\omega^\star_{n,ji},\ \begin{array}{c} j \in \mathcal{N}_{-i} \\ i \in \Gamma_{n,j}(K) \end{array}\right\} \tag{1}$$

as we take into account the possibility that node i may have been paired to some other node j. As mentioned earlier, if two nodes, say i and j, are within wireless range of each other, then they can establish a secure link if at least one of the events $i \in \Gamma_{n,j}(K)$ or $j \in \Gamma_{n,i}(K)$ takes place. Note that both events can take place, in which case the memory modules of node i and j each contain the distinct keys $\omega^\star_{n,ij}$ and $\omega^\star_{n,ji}$. By construction this scheme supports node-to-node authentication.

2.2. Gradual deployment

Initially n distinct node identities were generated, one for each of the n nodes, and the key rings

$$\Sigma_{n,1}(K), \ldots, \Sigma_{n,n}(K)$$

were constructed offline as indicated above—here n stands for the maximum possible network size and should be selected large enough. This key selection procedure does not require the physical presence of the sensor entities and can be implemented completely at the software level. We now describe how this offline pairwise key distribution scheme can accommodate gradual network deployment in consecutive stages.

In the initial phase of deployment, with $0 < \gamma_1 < 1$, let $\lfloor \gamma_1 n \rfloor$ sensors be produced and given the labels $1, \ldots, \lfloor \gamma_1 n \rfloor$. The key rings

$$\Sigma_{n,1}(K), \ldots, \Sigma_{n,\lfloor \gamma_1 n \rfloor}(K)$$

are then inserted into the memory modules of the sensors $1, \ldots, \lfloor \gamma_1 n \rfloor$, respectively. Imagine now that more sensors are needed, say $\lfloor \gamma_2 n \rfloor - \lfloor \gamma_1 n \rfloor$ sensors with $0 < \gamma_1 < \gamma_2 \le 1$. Then, $\lfloor \gamma_2 n \rfloor - \lfloor \gamma_1 n \rfloor$ additional sensors would be produced, this second batch of sensors would be assigned labels $\lfloor \gamma_1 n \rfloor + 1, \ldots, \lfloor \gamma_2 n \rfloor$, and the key rings

$$\Sigma_{n,\lfloor \gamma_1 n \rfloor+1}(K), \ldots, \Sigma_{n,\lfloor \gamma_2 n \rfloor}(K)$$

would be inserted into their memory modules. Once this is done, these $\lfloor \gamma_2 n \rfloor - \lfloor \gamma_1 n \rfloor$ new sensors are added to the network (which now comprises $\lfloor \gamma_2 n \rfloor$ deployed sensors). This step may be repeated a number of times: For some finite integer ℓ, consider positive scalars $0 < \gamma_1 < \cdots < \gamma_\ell \le 1$ (with $\gamma_0 = 0$ by convention). We can then deploy the sensor network in ℓ consecutive phases, with the kth phase adding $\lfloor \gamma_k n \rfloor - \lfloor \gamma_{k-1} n \rfloor$ new nodes to the network for each $k = 1, \ldots, \ell$.

2.3. Earlier results for the single phase case

The pairwise distribution scheme naturally gives rise to the following class of well-structured random graphs: With $n = 2, 3, \ldots$ and a positive integer K such that $K < n$, 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

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

Let $\mathbb{H}(n; K)$ denote the undirected random graph on the vertex set $\{1, \ldots, n\}$ induced by the adjacency notion (2). The following zero–one law for connectivity is established in [14,15] and is given here for easy reference.

Theorem 2.1. *For each positive integer K, it holds that*

$$\lim_{n\to\infty} \mathbb{P}\left[\mathbb{H}(n;K) \text{ is connected}\right] = \begin{cases} 0 & \text{if } K = 1 \\ 1 & \text{if } K \ge 2. \end{cases} \tag{3}$$

Moreover, for any $K \ge 2$, we have

$$\mathbb{P}\left[\mathbb{H}(n;K) \text{ is connected}\right] \ge 1 - \frac{155}{n^3}, \qquad \begin{matrix} n \ge 16, \\ K = 2, \ldots, n-1. \end{matrix} \tag{4}$$

The random graph $\mathbb{H}(n; K)$ is also known in the literature on random graphs as the K-out-n random graph. The first part of Theorem 2.1 can be derived from an earlier result on vertex connectivity obtained by Fenner and Frieze [20, Theorem 2.1, p. 348].

Some of the conclusions given earlier (at the end of Section 1.3) relied on the following monotonicity: Fix $n = 1, 2, \ldots$. An easy coupling argument shows that for positive integers K and K' such that $K < K' < n$, we have

$$\mathbb{P}\left[\mathbb{H}(n;K) \text{ is connected}\right] \le \mathbb{P}\left[\mathbb{H}(n;K') \text{ is connected}\right] \tag{5}$$

as expected.

3. On gradual deployment

With the network deployed gradually over time as described in Section 2.2, we seek to understand how the parameter K needs to be scaled with large n to ensure that connectivity is *maintained* a.a.s. throughout gradual deployment. The following terminology will be useful in what follows: A scaling is any mapping $K : \mathbb{N}_0 \to \mathbb{N}_0$ such that

$$K_n < n, \quad n = 2, 3, \dots .$$

Consider positive integers $n = 2, 3, \dots$ and K with $K < n$. With γ in the interval $(0, 1)$, let $\mathbb{H}_\gamma(n; K)$ denote the subgraph of $\mathbb{H}(n; K)$ restricted to the set of nodes $\{1, \dots, \lfloor \gamma n \rfloor\}$. The fact that $\mathbb{H}(n; K)$ is connected does *not* imply that $\mathbb{H}_\gamma(n; K)$ is automatically connected. Indeed, with distinct nodes $i, j = 1, \dots, \lfloor \gamma n \rfloor$, the path that exists in $\mathbb{H}(n; K)$ between nodes i and j (as a result of the assumed connectivity of $\mathbb{H}(n; K)$) may comprise edges that are *not* in $\mathbb{H}_\gamma(n; K)$. We write

$$P_\gamma(n; K) := \mathbb{P}\left[\, \mathbb{H}_\gamma(n; K) \text{ is connected} \,\right].$$

We shall also use the notation

$$P^\star_\gamma(n; K) := \mathbb{P}\left[\, \mathbb{H}_\gamma(n; K) \text{ contains no isolated nodes} \,\right].$$

3.1. Basic zero–one laws

The next result constitutes an analog of Theorem 2.1 in this new setting, and shows that gradual deployment has a significant impact on the dimensioning of the pairwise scheme.

Theorem 3.1. *With γ in the unit interval $(0, 1)$ and $c > 0$, consider a scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$ such that*

$$K_n \sim \frac{c}{\gamma} \log n. \tag{6}$$

Then, we have $\lim_{n\to\infty} P_\gamma(n; K_n) = 1$ whenever $c > 1$.

The random graphs $\mathbb{H}(n; K)$ and $\mathbb{H}_\gamma(n; K)$ have very different neighborhood structures. Indeed, any node in $\mathbb{H}(n; K)$ has degree at least K, so that no node is ever isolated in $\mathbb{H}(n; K)$. However, there is a positive probability that isolated nodes exist in $\mathbb{H}_\gamma(n; K)$. In fact, we have the following zero–one law.

Theorem 3.2. *Fix γ in the unit interval $(0, 1)$. For any scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$ such that (6) holds for some $c > 0$, we have*

$$\lim_{n\to\infty} P^\star_\gamma(n; K_n) = \begin{cases} 0 & \text{if } c < r(\gamma) \\ 1 & \text{if } r(\gamma) < c \end{cases} \tag{7}$$

where the threshold $r(\gamma)$ is given by

$$r(\gamma) := \left(1 - \frac{\log(1-\gamma)}{\gamma}\right)^{-1}. \tag{8}$$

It is easy to check that $r(\gamma)$ is a decreasing function of γ on the interval $[0, 1]$ with $\lim_{\gamma\downarrow 0} r(\gamma) = \frac{1}{2}$ and $\lim_{\gamma\uparrow 1} r(\gamma) = 0$. Since a connected graph has no isolated nodes, Theorem 3.2 yields $\lim_{n\to\infty} P_\gamma(n; K_n) = 0$ if the scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$ satisfies (6) with $c < r(\gamma)$. The following corollary is immediate from Theorem 3.1.

Corollary 3.3. *Fix γ in the unit interval $(0, 1)$. For any scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$ such that (6) holds for some $c > 0$, we have*

$$\lim_{n\to\infty} P_\gamma(n; K_n) = \begin{cases} 0 & \text{if } c < r(\gamma) \\ 1 & \text{if } 1 < c \end{cases} \tag{9}$$

with $r(\gamma)$ given by (8).

Corollary 3.3 does not provide a full zero–one law for the connectivity of $\mathbb{H}_\gamma(n; K_n)$ as there is a gap between the threshold $r(\gamma)$ of the zero-law and the threshold 1 of the one-law. Admittedly the gap between the thresholds of the zero-law and the one-law is quite small with $\frac{1}{2} < 1 - r(\gamma) < 1$. However, the simulation results in Section 5 strongly suggest the existence of a full zero–one law for $P_\gamma(n; K_n)$ with a threshold close to $r(\gamma)$. As a result, we conjecture that the analog of Theorem 3.2 also holds for graph connectivity. This is not too far-fetched since for many classes of random graphs, the absence of isolated nodes and graph connectivity are known to be asymptotically equivalent properties, e.g., Erdős–Rényi graphs [21], geometric random graphs [22] and random key graphs [7,9], among others.

3.2. Continuous connectivity throughout consecutive phases

Finally, we turn to gradual network deployment in several phases as discussed in Section 2.2. Given the scalars $0 < \gamma_1 < \dots < \gamma_\ell \leq 1$, we give conditions on how to scale K as a function of n such that $\mathbb{H}_{\gamma_k}(n; K)$ is a.a.s. connected for each

$k = 1, 2, \ldots, \ell$. With γ in $(0, 1]$, we introduce the event

$$C_{n,\gamma}(K) := \left[\, \mathbb{H}_\gamma(n; K) \text{ is connected} \,\right].$$

Theorem 3.4. *With $0 < \gamma_1 < \gamma_2 < \cdots < \gamma_\ell \le 1$, consider a scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$ such that*

$$K_n \sim \frac{c}{\gamma_1} \log n \tag{10}$$

for some $c > 1$. Then, we have

$$\lim_{n\to\infty} \mathbb{P}\left[\, C_{n,\gamma_1}(K_n) \cap \cdots \cap C_{n,\gamma_\ell}(K_n) \,\right] = 1. \tag{11}$$

The event $C_{\gamma_1,n}(K_n) \cap \cdots \cap C_{\gamma_\ell,n}(K_n)$ corresponds to the network being connected in *each* of the ℓ phases of deployment—in other words, on that event the sensors do form a connected network at each phase of the gradual deployment. Theorem 3.4 shows that the condition (10) (with $c > 1$) is enough to ensure that the network remains a.a.s. connected as more sensors are deployed over time.

Proof. With the notation in the statement of Theorem 3.4, it is plain that (11) will hold, provided

$$\lim_{n\to\infty} \mathbb{P}\left[C_{n,\gamma_k}(K_n)\right] = 1, \quad k = 1, \ldots, \ell. \tag{12}$$

For each $k = 1, 2, \ldots, \ell$, we note that

$$\frac{c}{\gamma_1} \log n = \frac{c_k}{\gamma_k} \log n \quad \text{with } c_k := c\, \frac{\gamma_k}{\gamma_1}$$

for all $n = 1, 2, \ldots$. But $c > 1$ implies $c_k > 1$ since $\gamma_1 < \cdots < \gamma_\ell$. As a result, $\mathbb{H}_{\gamma_k}(n; K_n)$ is a.a.s. connected by virtue of Theorem 3.1 applied to $\mathbb{H}_{\gamma_k}(n; K)$, and (12) indeed holds. □

4. On key ring sizes

According to Theorem 2.1, very small values of K suffice for a.a.s. connectivity of the random graph $\mathbb{H}(n; K)$. However the mere fact that $\mathbb{H}(n; K)$ becomes connected even with very small values of K does not imply that the *number* of keys needed to achieve connectivity is necessarily small—this is because the pairwise scheme generates key rings whose sizes are variable over a wide range. In this section, we explore how this variability in key ring sizes behaves as the number of nodes becomes large.

4.1. A concentration result

Fix $n = 2, 3, \ldots$ and a positive integer K with $K < n$. For each $i = 1, 2, \ldots, n$, the key ring $\Sigma_{n,i}(K)$ assigned to node i has size

$$\begin{aligned} |\Sigma_{n,i}(K)| &= |\Gamma_{n,i}(K)| + \sum_{j=1,\quad j\neq i}^{n} \mathbf{1}\left[i \in \Gamma_{n,j}(K)\right] \\ &= K + \sum_{j=1,\, j\neq i}^{n} \mathbf{1}\left[i \in \Gamma_{n,j}(K)\right] \end{aligned} \tag{13}$$

since $|\Gamma_{n,i}(K)| = K$ by construction. This follows from the definition (1) and shows that the size of any key ring is random with a range from K to $K + (n-1)$. Yet, it is nevertheless the case that

$$\mathbb{E}\left[|\Sigma_{n,i}(K)|\right] = K + (n-1)\frac{K}{n-1} = 2K \tag{14}$$

since

$$\mathbb{P}\left[i \in \Gamma_{n,j}(K)\right] = \frac{\binom{n-2}{K-1}}{\binom{n-1}{K}} = \frac{K}{n-1}, \quad j \in \mathcal{N}_{-i}. \tag{15}$$

This suggests that perhaps the size of key rings assumes values much larger than the mean with small probability.

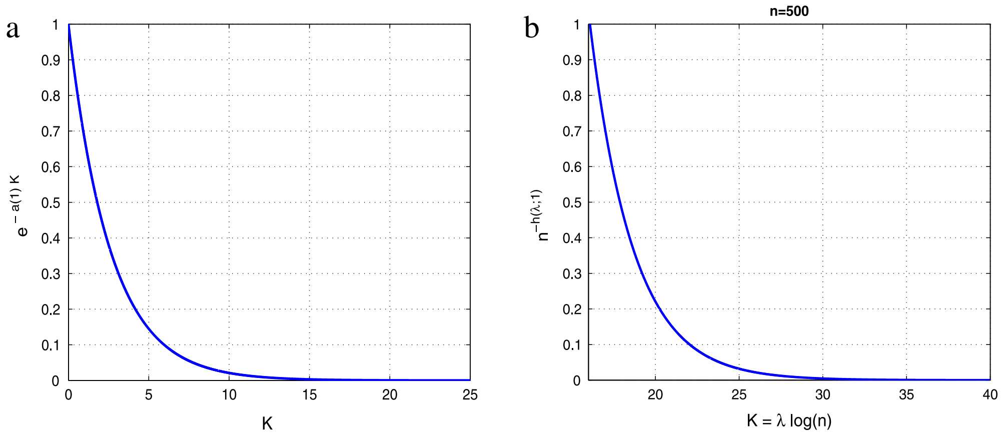

Fig. 1. (a) The upper bound on the probability that $\Sigma_{n,1}(K_n) \geq 3K_n$ as given by Lemma 4.1 (with $c = 1$). (b) The upper bound on the probability that $M_n(K_n) \geq 3K_n$ as given by Theorem 4.2 (when $c = 1$). We have taken $K_n \sim \lambda \log n$ with $n = 500$ and vary λ between $\lambda^\star$ and $3\lambda^\star$.

Our first result shows that the size of key rings indeed tends to concentrate around their common expected value; this is a consequence of standard concentration bounds for binomial rvs. To state the result, set

$$a(\tau) := (1+\tau) \cdot \log(1+\tau) - \tau, \quad \tau > -1, \tag{16}$$

and

$$b(\tau) := \begin{cases} 2 & \text{if } 0 < \tau < 1 \\ 1 & \text{if } 1 \leq \tau. \end{cases} \tag{17}$$

Lemma 4.1. *Consider positive integers $n = 2, 3, \ldots$ and K with $K < n$. For any $c > 0$, we have*

$$\mathbb{P}\left[\ \big||\Sigma_{n,1}(K)| - 2K\big| > cK\ \right] \leq b(c) \cdot e^{-Ka(c)} \tag{18}$$

with $a(c) > 0$ and $b(c)$ given by (16) and (17), respectively.

Lemma 4.1, which is established in Section 8.2, has several consequences: If the parameter K is scaled according to some scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$ such that $\lim_{n\to\infty} K_n = \infty$, then as an easy consequence of Lemma 4.1 we have

$$\frac{|\Sigma_{n,1}(K_n)|}{2K_n} \xrightarrow{P}_n 1$$

as soon as $\lim_{n\to\infty} K_n = \infty$. Thus, when K_n becomes large with n, although the key ring size $|\Sigma_{n,1}(K_n)|$ could in principle fluctuate from K_n to $K_n + (n-1)$, it has a propensity to hover about its mean $2K_n$.

It also follows from Lemma 4.1 that in the large n limit, the number of keys $|\Sigma_{n,1}(K_n)|$ stored at a node will be between $(2-c)K_n$ and $(2+c)K_n$ with high probability for any $c > 0$. For the sake of concreteness consider the case $c = 1$ (so that $b(1) = 1$): Since we always have $|\Sigma_{n,1}(K_n)| \geq K_n$ (e.g., see (13)), the inequality (18) reduces to

$$\mathbb{P}\left[\ |\Sigma_{n,1}(K_n)| > 3K_n\ \right] \leq e^{-a(1)K_n} \tag{19}$$

for all $n = 2, 3, \ldots$. To see how fast $\mathbb{P}\left[|\Sigma_{n,1}(K_n)| > 3K_n\right]$ decays to zero, consider Fig. 1(a) where the upper bound appearing at (19) is shown as a function of K_n—note that $\mathbb{P}\left[|\Sigma_{n,1}(K_n)| > 3K_n\right]$ is already negligible (i.e., "vanishes" for all practical purposes) for $K_n \geq 15$.

4.2. Maximal key ring sizes when scaling K

In order to address worst case situations, we consider the maximal key ring size defined by

$$M_n(K) := \max_{i=1,\ldots,n} |\Sigma_{n,i}(K)|. \tag{20}$$

Since every key appears in exactly two different key rings, the relation

$$\sum_{i=1}^{n} |\Sigma_{n,i}(K)| = 2nK$$

holds by construction, whence $M_n(K) \geq 2K$. A result analogous to (19) also holds for the maximal key ring size when attention is restricted to the subclass of logarithmic scalings.

Theorem 4.2. *Consider a scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$ such that*

$$K_n \sim \lambda \log n \tag{21}$$

for some $\lambda > 0$. Then, for any $c > 0$, the bound

$$\mathbb{P}\left[\, M_n(K_n) > (2+c)K_n \,\right] \le n^{-(h(\lambda;c)+o(1))} \tag{22}$$

holds for all $n = 1, 2, \ldots$ with

$$h(\lambda; c) := -1 + \lambda a(c), \quad c > 0. \tag{23}$$

Moreover, there exists $c(\lambda) > 0$ such that $c \to h(\lambda; c)$ changes sign exactly once on $\mathbb{R}_+$ at $c(\lambda)$ with

$$h(\lambda; c) > 0, \quad c(\lambda) < c. \tag{24}$$

Theorem 4.2 is established in Section 8.3. From the discussion given there it is easy to check that the mapping $\lambda \to c(\lambda)$ is strictly decreasing on $(0, \infty)$ with $1 < c(\lambda)$ (resp. $c(\lambda) < 1$) if and only if $\lambda < \lambda^\star$ (resp. $\lambda^\star < \lambda$) where the parameter value $\lambda^\star$ is given by

$$\lambda^\star := (2\log 2 - 1)^{-1} \simeq 2.6. \tag{25}$$

Theorem 4.2 shows that under the condition (21) (with some $\lambda > 0$) on the scaling, with high probability in the large n limit, the maximum key ring size $M_n(K_n)$ will not exceed $(2+c)K_n$ whenever $c(\lambda) < c$. For instance, for $c = 1$ the maximal key ring size is now seen to be less than $3K_n$ with high probability *provided* λ has been selected so that $c(\lambda) < 1$ (in which case $h(\lambda; 1) > 0$)—in view of earlier comments, this happens if and only if $\lambda > \lambda^\star \simeq 2.6$.

To get a better sense as to how fast $\mathbb{P}[M_n(K_n) > 3K_n]$ decays to zero, we have plotted the upper bound appearing at (22) for $c = 1$, $n = 500$ and $K = \lambda \log n$ as λ varies (or equivalently, as K varies). The basic upper bound (69) in the proof of Theorem 4.2 shows that if $K_n = \lambda \log n$, then the upper bound appearing at (22) coincides with $n^{-h(\lambda;c)}$.[1] In this setting, Fig. 1(b) depicts $n^{-h(\lambda;1)}$ (implicitly as a function of K) as λ varies between $\lambda^\star$ and $3\lambda^\star$. This time the upper bound becomes negligible for $K > 30$.

We can combine these results with earlier results concerning the secure connectivity of the pairwise scheme when the network is deployed gradually over time—it should be clear from Theorem 3.4 why scalings of the form (21) are of particular interest here. Combining Theorems 3.1 and 4.2 allows us to reach the following useful conclusions: With $0 < \gamma_1 < \gamma_2 < \cdots < \gamma_\ell \le 1$, consider a scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$ such that $K_n = \Theta(\log n)$ with

$$K_n \ge \max\left(\lambda^\star, \frac{1}{\gamma_1}\right) \cdot \log n$$

for all n sufficiently large. The following statements then hold true: (i) The maximum number of keys kept in the memory module of each sensor will be a.a.s. less than $3K_n$; (ii) The network deployed gradually in ℓ steps (as in Section 2) will be a.a.s. connected in each of the ℓ phases of deployment. Thus, key rings of size $\Theta(\log n)$ are sufficient to ensure secure connectivity under the pairwise scheme (even when the network is gradually deployed).

4.3. Maximal key ring sizes with fixed K

We close with a concentration result for the maximal key ring size when the parameter K is not scaled with n; in light of Theorem 2.1, this situation may be of interest when the network is deployed in a single phase: A simple coupling argument shows that for each $n = 1, 2, \ldots$ and positive integers K and K', we have[2] $M_n(K) \le_{st} M_n(K')$ whenever $K < K' < n$, i.e.,

$$\mathbb{P}\left[\, M_n(K) > t \,\right] \le \mathbb{P}\left[M_n(K') > t\right], \quad t \ge 0. \tag{26}$$

Invoking Theorem 4.2 with $c > 0$ and $\lambda > 0$, we conclude from (26) that

$$\mathbb{P}\left[\, M_n(K) > (2+c)\lceil \lambda \log n \rceil \,\right] \le n^{-(h(\lambda;c)+o(1))} \tag{27}$$

for all n sufficiently large as soon as $K < \lceil \lambda \log n \rceil < n$. Thus, whenever we have $c(\lambda) < c$, the maximal key ring size $M_n(K)$ will be on the order $O(\log n)$ with very high probability when n is large. The next result shows that sharper bounds are in fact available.

[1] The term $o(1)$ in (22) accounts for the fact that $\lambda \log n$ may not be an integer.

[2] The notation $\le_{st}$ stands for less or equal in the strong stochastic ordering.

Please cite this article in press as: O. Yağan, A.M. Makowski, On the scalability of the random pairwise key predistribution scheme: Gradual deployment and key ring sizes, Performance Evaluation (2013), http://dx.doi.org/10.1016/j.peva.2013.03.001

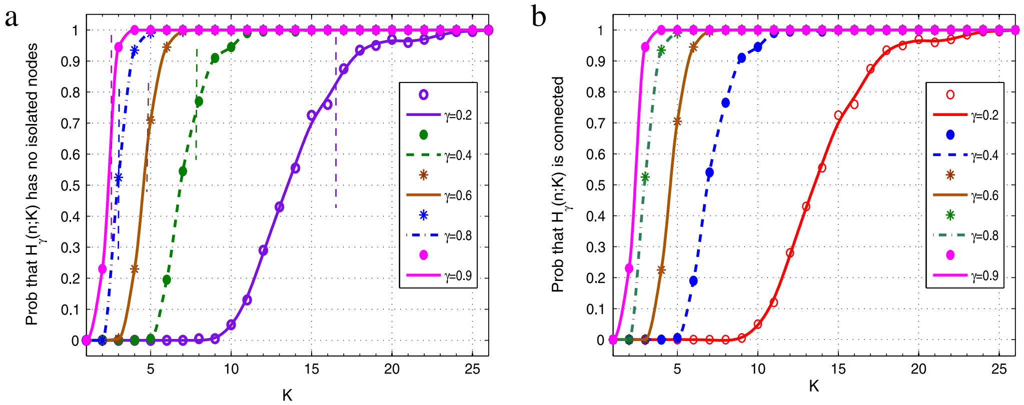

Fig. 2. (a) Probability that $\mathbb{H}_\gamma(n; K)$ contains no isolated nodes for $n = 1000$ obtained through 200 experiments. Vertical dashed lines stand for the critical thresholds asserted by Theorem 3.2. The theoretical findings are in perfect agreement with the simulation results. (b) Probability that $\mathbb{H}_\gamma(n; K)$ is connected for $n = 1000$ obtained through 200 experiments. The curves are almost indistinguishable from the corresponding ones of Fig. 2(a); this supports the claim that the absence of isolated nodes and connectivity are asymptotically equivalent properties.

Theorem 4.3. *For each positive integer K and $c > 0$ we have*

$$\mathbb{P}\left[M_n(K) > K + c \cdot \frac{\log n}{\log \log n}\right] \leq n^{1-c+o(1)}. \tag{28}$$

Theorem 4.3 is established in Section 8.4 and states that for large n, the maximal key ring size is on the order $O(\frac{\log n}{\log \log n})$ with very high probability whenever K is a fixed integer. Indeed we note from (28) that

$$\lim_{n\to\infty} \mathbb{P}\left[M_n(K) > K + c \cdot \frac{\log n}{\log \log n}\right] = 0, \quad c > 1.$$

The leading exponents of n appearing in the bounds at (27) and (28) can be easily compared, with $-h(\lambda; c) < 1 - c$ if and only if $c < \lambda a(c)$. This could be used to explore the relative rate at which the relevant probabilities are driven to zero.

5. Limited simulation experiments

Before moving on to the proofs of the results presented in Sections 3 and 4, we pause to present experimental results in support of the theoretical findings discussed earlier.

5.1. Gradual deployment

In each set of experiments, we fix n and γ. Then, we generate random graphs $\mathbb{H}_\gamma(n; K)$ for each $K = 1, \dots, K_{\max}$ where the maximal value $K_{\max}$ is selected large enough. In each case, we check whether the generated graph has isolated nodes and is connected. We repeat the process 200 times for each pair of values γ and K in order to estimate the probabilities of the events of interest. For various values of γ, Fig. 2(a) depicts the estimated probability $P^\star_\gamma(n; K)$ that $\mathbb{H}_\gamma(n; K)$ contains no isolated nodes as a function of K. Here, n is taken to be 1000. The plots in Fig. 2(a) clearly confirm the claims of Theorem 3.2: In each case $P^\star_\gamma(n; K)$ exhibits a threshold behavior as K increases, and the transitions from $P^\star_\gamma(n; K) = 0$ to $P^\star_\gamma(n; K) = 1$ take place around $K_n(\gamma) = r(\gamma)\frac{\log n}{\gamma}$ as dictated by Theorem 3.2; the critical value $K_n(\gamma)$ is shown by a vertical dashed line in each plot.

Similarly, Fig. 2(b) shows the estimated probability $P_\gamma(n; K)$ as a function of K for various values of γ with $n = 1000$. For each specified γ, we see that the variation of $P_\gamma(n; K)$ with K is almost indistinguishable from that of $P^\star_\gamma(n; K)$ supporting the claim that $P_\gamma(n; K)$ exhibits a full zero–one law similar to that of Theorem 3.2 with a threshold behaving like $r(\gamma)$. We can also conclude by monotonicity that $P_\gamma(n; K) = 1$ whenever (6) holds with $c > 1$; this verifies Theorem 3.1. Furthermore, it is evident from Fig. 2(b) that for given K and n, $P_\gamma(n; K)$ increases as γ increases supporting Theorem 3.4.

5.2. Key ring sizes

Next, we present simulation results that illustrate Lemma 4.1 and Theorem 4.2: For fixed values of n and K we have constructed key rings according to the mechanism presented in Section 2. For each pair of parameters n and K, the experiments have been repeated 1000 times yielding $1000 \times n$ key rings for each parameter pair. The results are depicted

Please cite this article in press as: O. Yağan, A.M. Makowski, On the scalability of the random pairwise key predistribution scheme: Gradual deployment and key ring sizes, Performance Evaluation (2013), http://dx.doi.org/10.1016/j.peva.2013.03.001

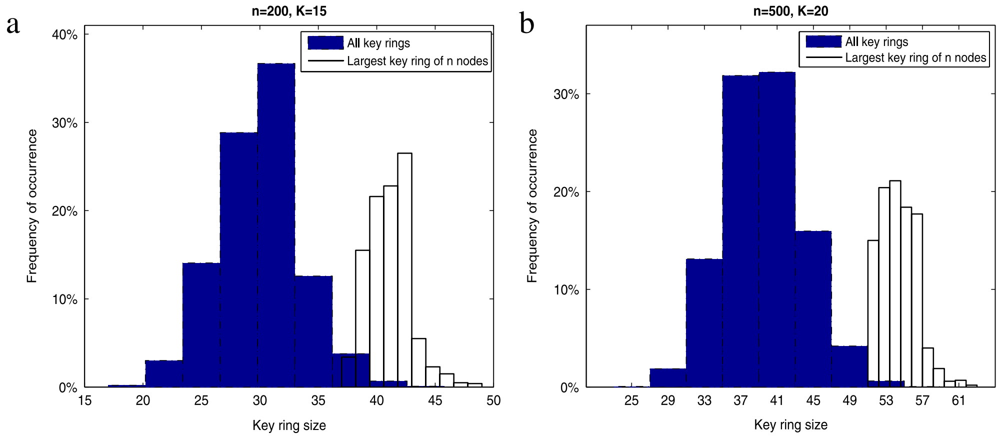

Fig. 3. Key ring sizes observed in 1000 experiments for (a) $n = 200$, $K = 15$, and for (b) $n = 500$, $K = 20$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

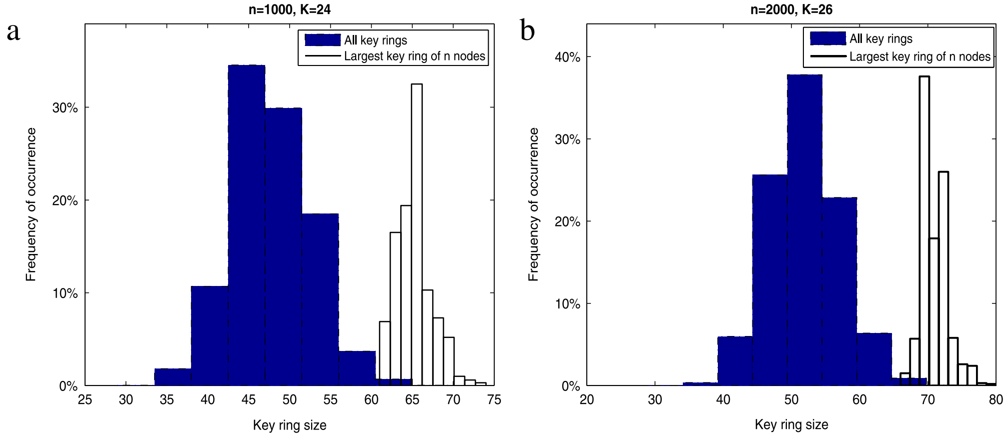

Fig. 4. Key ring sizes observed in 1000 experiments for (a) $n = 1000$, $K = 24$, and for (b) $n = 2000$, $K = 26$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

in Figs. 3(a)–4(b) which show the key ring sizes according to their frequency of occurrence. The histograms filled with color (blue) correspond to the key ring $\Sigma_{n,1}(K)$, while the histograms that have colorless bins stand for the maximal key ring size $M_n(K)$, i.e., only the largest key ring amongst n nodes in an experiment.

It is immediate from Figs. 3(a)–4(b) that $|\Sigma_{n,1}(K)|$ tends to concentrate around $2K$, in agreement with the claim of Lemma 4.1. As would be expected, this concentration becomes more evident as n gets large. It is also clear that, in almost all cases the maximum size of a key ring (over the n nodes) is less than $3K$, validating the claim of Theorem 4.2. In particular, the largest key rings observed are 49, 64, 74, 80 for the cases ($n = 200$, $K = 15$), ($n = 500$, $K = 20$), ($n = 1000$, $K = 24$) and ($n = 2000$, $K = 26$), respectively.

6. A proof of Theorem 3.1

Consider γ in the interval $(0, 1)$, and fix $n = 2, 3, \ldots$ and a positive integer $K \geq 2$. Throughout the discussion, in order to avoid degenerate situations, we take n sufficiently large so that the conditions

$$2(K+1) < n, \qquad K+1 \leq n - \lfloor \gamma n \rfloor \quad \text{and} \quad 2 < \gamma n \tag{29}$$

are all enforced; this has no bearing on the final result since we eventually let n go to infinity under a scaling which satisfies (6).

For any non-empty subset R contained in $\{1, \ldots, \lfloor \gamma n \rfloor\}$, we say that R is *isolated* in $\mathbb{H}_\gamma(n; K)$ if there are no edges (in $\mathbb{H}_\gamma(n; K)$) between the nodes in R and the nodes in its complement $R^{c|\gamma} := \{1, \ldots, \lfloor \gamma n \rfloor\} - R$. This is characterized by the event $B_{n,\gamma}(K; R)$ given by

$$B_{n,\gamma}(K; R) := \left[i \notin \Gamma_{n,j}(K), j \notin \Gamma_{n,i}(K), \ i \in R, \ j \in R^{c|\gamma} \right].$$

The discussion starts with the following basic observation (already used in deriving the one-law for connectivity in Erdős–Rényi graphs, e.g., see [23, Section 3.4.2, p. 42]): If $\mathbb{H}_\gamma(n;K)$ is *not* connected, then there must exist a non-empty subset R of nodes contained in $\{1,\ldots,\lfloor \gamma n\rfloor\}$, and isolated in $\mathbb{H}_\gamma(n;K)$. This is captured by the inclusion

$$C_{n,\gamma}(K)^c \subseteq \cup_{R\in\mathcal{N}_{n,\gamma}} \; B_{n,\gamma}(K;R) \tag{30}$$

with $\mathcal{N}_{n,\gamma}$ denoting the collection of all non-empty subsets of $\{1,\ldots,\lfloor\gamma n\rfloor\}$. This union needs only be taken over all non-empty subsets R of $\{1,\ldots,\lfloor\gamma n\rfloor\}$ with $1 \le |R| \le \lfloor \frac{\lfloor\gamma n\rfloor}{2}\rfloor$, and for future use it is useful to note that $\lfloor\frac{\lfloor\gamma n\rfloor}{2}\rfloor = \lfloor\frac{\gamma n}{2}\rfloor$. A standard union bound argument immediately gives

$$\begin{aligned}\mathbb{P}\left[C_{n,\gamma}(K)^c\right] &\le \sum_{R\in\mathcal{N}_{n,\gamma}:\, 1\le|R|\le\lfloor\frac{\gamma n}{2}\rfloor} \mathbb{P}\left[B_{n,\gamma}(K;R)\right]\\ &= \sum_{r=1}^{\lfloor\frac{\gamma n}{2}\rfloor}\left(\sum_{R\in\mathcal{N}_{n,\gamma,r}} \mathbb{P}\left[B_{n,\gamma}(K;R)\right]\right)\end{aligned} \tag{31}$$

where $\mathcal{N}_{n,\gamma,r}$ denotes the collection of all subsets of $\{1,\ldots,\lfloor\gamma n\rfloor\}$ with exactly r elements.

For each $r = 1,\ldots,\lfloor\gamma n\rfloor$, when $R=\{1,\ldots,r\}$ we simplify the notation by writing $B_{n,\gamma,r}(K) := B_{n,\gamma}(K;R)$. Under the enforced assumptions, we have

$$\mathbb{P}\left[B_{n,\gamma}(K;R)\right] = \mathbb{P}\left[B_{n,\gamma,r}(K)\right], \quad R\in\mathcal{N}_{n.\gamma,r}$$

by exchangeability, and the expression

$$\sum_{R\in\mathcal{N}_{\gamma,r}} \mathbb{P}\left[B_{n,\gamma}(K;R)\right] = \binom{\lfloor\gamma n\rfloor}{r}\,\mathbb{P}\left[B_{n,\gamma,r}(K)\right]$$

follows since $|\mathcal{N}_{n,\gamma,r}| = \binom{\lfloor\gamma n\rfloor}{r}$. Substituting into (31) yields the bounds

$$\mathbb{P}\left[C_{n,\gamma}(K)^c\right] \le \sum_{r=1}^{\lfloor\frac{\gamma n}{2}\rfloor}\binom{\lfloor\gamma n\rfloor}{r}\,\mathbb{P}\left[B_{n,\gamma,r}(K)\right]. \tag{32}$$

Under the enforced assumptions, we get

$$\mathbb{P}\left[B_{n,\gamma,r}(K)\right] = \left(\frac{\binom{n-\lfloor\gamma n\rfloor+r-1}{K}}{\binom{n-1}{K}}\right)^r\left(\frac{\binom{n-r-1}{K}}{\binom{n-1}{K}}\right)^{\lfloor\gamma n\rfloor - r}. \tag{33}$$

To see why this last relation holds, recall that for the set $\{1,\ldots,r\}$ to be isolated *in* $\mathbb{H}_\gamma(n;K)$ we need that (i) each of the nodes $r+1,\ldots,\lfloor\gamma n\rfloor$ are adjacent only to nodes *outside* the set of nodes $\{1,\ldots,r\}$; and (ii) none of the nodes $1,\ldots,r$ are adjacent with any of the nodes $r+1,\ldots,\lfloor\gamma n\rfloor$—this last requirement does not preclude adjacency with any of the nodes $\lfloor\gamma n\rfloor+1,\ldots,n$. Substituting (33) into (32), we conclude that

$$\mathbb{P}\left[C_{n,\gamma}(K)^c\right] \le \sum_{r=1}^{\lfloor\frac{\gamma n}{2}\rfloor}\binom{\lfloor\gamma n\rfloor}{r}\left(\frac{\binom{n-\lfloor\gamma n\rfloor+r-1}{K}}{\binom{n-1}{K}}\right)^r\cdot\left(\frac{\binom{n-r-1}{K}}{\binom{n-1}{K}}\right)^{\lfloor\gamma n\rfloor - r} \tag{34}$$

with conditions (29) ensuring that the binomial coefficients are all well defined.

The remainder of the proof consists in bounding each of the terms in (34) with the help of several standard bounds. First we recall the well-known bound

$$\binom{\lfloor\gamma n\rfloor}{r} \le \left(\frac{\lfloor\gamma n\rfloor e}{r}\right)^r, \quad r = 1,\ldots,\lfloor\gamma n\rfloor. \tag{35}$$

Next, for $0\le K\le x\le y$, we note that

$$\frac{\binom{x}{K}}{\binom{y}{K}} = \prod_{\ell=0}^{K-1}\left(\frac{x-\ell}{y-\ell}\right) \le \left(\frac{x}{y}\right)^K$$

since $\frac{x-\ell}{y-\ell}$ decreases as ℓ increases from $\ell=0$ to $\ell=K-1$.

Please cite this article in press as: O. Yağan, A.M. Makowski, On the scalability of the random pairwise key predistribution scheme: Gradual deployment and key ring sizes, Performance Evaluation (2013), http://dx.doi.org/10.1016/j.peva.2013.03.001

Now pick $r = 1, \ldots, \lfloor \gamma n \rfloor$. Under (29) we can apply these bounds to obtain

$$\begin{aligned}\binom{\lfloor \gamma n\rfloor}{r}\left(\frac{\binom{n-\lfloor \gamma n\rfloor+r-1}{K}}{\binom{n-1}{K}}\right)^{r}\cdot\left(\frac{\binom{n-r-1}{K}}{\binom{n-1}{K}}\right)^{\lfloor \gamma n\rfloor-r} &\leq \left(\frac{\lfloor \gamma n\rfloor e}{r}\right)^{r}\cdot\left(\frac{n-\lfloor \gamma n\rfloor+r-1}{n-1}\right)^{rK}\cdot\left(\frac{n-r-1}{n-1}\right)^{K(\lfloor \gamma n\rfloor-r)}\\ &\leq \left(\frac{\gamma ne}{r}\right)^{r}\left(1-\frac{\lfloor \gamma n\rfloor-r}{n-1}\right)^{rK}\left(1-\frac{r}{n-1}\right)^{K(\lfloor \gamma n\rfloor-r)}\\ &\leq (\gamma ne)^{r}\cdot\left(1-\frac{\lfloor \gamma n\rfloor-r}{n}\right)^{rK}\cdot\left(1-\frac{r}{n}\right)^{K(\lfloor \gamma n\rfloor-r)}\\ &\leq (\gamma ne)^{r}\cdot e^{-\left(\frac{\lfloor \gamma n\rfloor-r}{n}\right)rK}\cdot e^{-\frac{r}{n}(\lfloor \gamma n\rfloor-r)K}.\end{aligned}$$

It is plain that

$$\begin{aligned}\mathbb{P}\left[C_{n,\gamma}(K)^{c}\right] &\leq \sum_{r=1}^{\lfloor \frac{\gamma n}{2}\rfloor}(\gamma ne)^{r}\cdot e^{-2\left(\frac{\lfloor \gamma n\rfloor-r}{n}\right)rK}\\ &\leq \sum_{r=1}^{\lfloor \frac{\gamma n}{2}\rfloor}\left(\gamma ne\cdot e^{-2\left(\frac{\lfloor \gamma n\rfloor-\lfloor \frac{\gamma n}{2}\rfloor}{n}\right)K}\right)^{r}. \qquad (36)\end{aligned}$$

Next, consider a scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$ such that (6) holds for some $c > 1$. This amounts to

$$K_n = \frac{c_n}{\gamma}\log n, \quad n = 1, 2, \ldots \qquad (37)$$

for some sequence $\mathbb{N}_0 \to \mathbb{R}_+ : n \to c_n$ such that $\lim_{n\to\infty} c_n = c$. Note that

$$a_n := \gamma ne\cdot e^{-2\left(\frac{\lfloor \gamma n\rfloor-\lfloor \frac{\gamma n}{2}\rfloor}{n}\right)K_n} = (\gamma e)\cdot n^{1-2c_n\left(\frac{\lfloor \gamma n\rfloor-\lfloor \frac{\gamma n}{2}\rfloor}{\gamma n}\right)}$$

for each $n = 1, 2, \ldots$. Since $\lim_{n\to\infty} c_n = c$, it is a simple matter to check that

$$\lim_{n\to\infty}\left(2c_n\left(\frac{\lfloor \gamma n\rfloor-\lfloor \frac{\gamma n}{2}\rfloor}{\gamma n}\right)\right) = c,$$

so that $\lim_{n\to\infty} a_n = 0$ by virtue of the fact that $c > 1$.

From (37) there exists a finite integer $n^\star$ such that the conditions

$$2(K_n + 1) < n, \qquad K_n + 1 \leq n - \lfloor \gamma n \rfloor \quad \text{and} \quad 2 < \gamma n \qquad (38)$$

hold for all $n \geq n^\star$. On that range, the bound (36) thus holds with K replaced by K_n, whence

$$\mathbb{P}\left[C_{n,\gamma}(K_n)^{c}\right] \leq \sum_{r=1}^{\lfloor \frac{\gamma n}{2}\rfloor}(a_n)^{r} \leq \sum_{r=1}^{\infty}(a_n)^{r} = \frac{a_n}{1-a_n} \qquad (39)$$

where for all n sufficiently large the summability of the geometric series is guaranteed by the convergence $\lim_{n\to\infty} a_n = 0$. This fact also yields the desired conclusion $\lim_{n\to\infty} \mathbb{P}\left[C_{n,\gamma}(K)^c\right] = 0$ via (39).

7. A proof of Theorem 3.2

Consider γ in $(0, 1)$. Moreover, fix $n = 2, 3, \ldots$ and a positive integer K such that $K < n$. For each $i = 1, \ldots, \lfloor \gamma n \rfloor$, we write

$$\chi_{n,\gamma,i}(K) := \mathbf{1}\left[\text{ Node } i \text{ is isolated in } \mathbb{H}_\gamma(n; K)\right].$$

The number of isolated nodes in $\mathbb{H}_\gamma(n; K)$ is simply given by

$$I_{n,\gamma}(K) := \sum_{i=1}^{\lfloor \gamma n\rfloor}\chi_{n,\gamma,i}(K),$$

and the random graph $\mathbb{H}_\gamma(n; K)$ has no isolated nodes if $I_{n,\gamma}(K) = 0$.

7.1. The method of first and second moments

We use the method of first moment [24, Eqn. (3.10), p. 55] and second moment [24, Remark 3.1, p. 55], an approach which relies on the bounds

$$1-\mathbb{E}\left[I_{n,\gamma}(K)\right] \le \mathbb{P}\left[I_{n,\gamma}(K)=0\right] \le 1-\frac{\left(\mathbb{E}\left[I_{n,\gamma}(K)\right]\right)^2}{\mathbb{E}\left[I_{n,\gamma}(K)^2\right]}. \tag{40}$$

First some computations: The rvs $\chi_{n,\gamma,1}(K),\ldots,\chi_{n,\gamma,\lfloor\gamma n\rfloor}(K)$ being exchangeable, we find

$$\mathbb{E}\left[I_{n,\gamma}(K)\right]=\lfloor\gamma n\rfloor\mathbb{E}\left[\chi_{n,\gamma,1}(K)\right] \tag{41}$$

and

$$\mathbb{E}\left[I_{n,\gamma}(K)^2\right]=\lfloor\gamma n\rfloor\mathbb{E}\left[\chi_{n,\gamma,1}(K)\right]+\lfloor\gamma n\rfloor(\lfloor\gamma n\rfloor-1)\mathbb{E}\left[\chi_{n,\gamma,1}(K)\chi_{n,\gamma,2}(K)\right] \tag{42}$$

by the binary nature of the rvs involved. It follows in the usual manner that

$$\frac{\mathbb{E}\left[I_{n,\gamma}(K)^2\right]}{\left(\mathbb{E}\left[I_{n,\gamma}(K)\right]\right)^2}=\frac{1}{\lfloor\gamma n\rfloor\mathbb{E}\left[\chi_{n,\gamma,1}(K)\right]}+\frac{\lfloor\gamma n\rfloor-1}{\lfloor\gamma n\rfloor}\frac{\mathbb{E}\left[\chi_{n,\gamma,1}(K)\chi_{n,\gamma,2}(K)\right]}{\left(\mathbb{E}\left[\chi_{n,\gamma,1}(K)\right]\right)^2}. \tag{43}$$

Pick a scaling $K:\mathbb{N}_0\to\mathbb{N}_0$ such that (6) holds for some $c>0$, and replace K by K_n in (40), so that

$$1-\mathbb{E}\left[I_{n,\gamma}(K_n)\right] \le \mathbb{P}\left[I_{n,\gamma}(K_n)=0\right] \le 1-\frac{\left(\mathbb{E}\left[I_{n,\gamma}(K_n)\right]\right)^2}{\mathbb{E}\left[I_{n,\gamma}(K_n)^2\right]} \tag{44}$$

for all $n=2,3,\ldots$. The next step relies on the two lemmas presented next; their proofs are available in Sections 7.3 and 7.4, respectively.

Lemma 7.1. *Consider γ in $(0,1)$ and a scaling $K:\mathbb{N}_0\to\mathbb{N}_0$ such that (6) holds for some $c>0$. We have*

$$\lim_{n\to\infty} n\mathbb{E}\left[\chi_{n,\gamma,1}(K_n)\right]=\begin{cases}0 & \text{if } r(\gamma)<c\\ \infty & \text{if } c<r(\gamma)\end{cases} \tag{45}$$

with $r(\gamma)$ specified by (8).

Lemma 7.2. *Consider γ in $(0,1)$ and a scaling $K:\mathbb{N}_0\to\mathbb{N}_0$ such that (6) holds for some $c>0$. We have*

$$\limsup_{n\to\infty}\left(\frac{\mathbb{E}\left[\chi_{n,\gamma,1}(K_n)\chi_{n,\gamma,2}(K_n)\right]}{\left(\mathbb{E}\left[\chi_{n,\gamma,1}(K_n)\right]\right)^2}\right)\le 1. \tag{46}$$

Assume $r(\gamma)<c$. Lemma 7.1 yields $\lim_{n\to\infty}\lfloor\gamma n\rfloor\mathbb{E}\left[\chi_{n,\gamma,1}(K_n)\right]=0$, whence

$$\lim_{n\to\infty}\mathbb{E}\left[I_{n,\gamma}(K_n)\right]=0.$$

Letting n go to infinity in the first inequality at (44) yields the one-law $\lim_{n\to\infty}\mathbb{P}\left[I_{n,\gamma}(K_n)=0\right]=1$.

Assume $c<r(\gamma)$. This time, Lemma 7.1 gives $\lim_{n\to\infty}\lfloor\gamma n\rfloor\mathbb{E}\left[\chi_{n,\gamma,1}(K_n)\right]=\infty$, and with the help of the relation (43), we find

$$\limsup_{n\to\infty}\frac{\mathbb{E}\left[I_{n,\gamma}(K_n)^2\right]}{\left(\mathbb{E}\left[I_{n,\gamma}(K_n)\right]\right)^2}\le 1$$

as we invoke Lemma 7.2. Inverting yields

$$1\le\liminf_{n\to\infty}\frac{\left(\mathbb{E}\left[I_{n,\gamma}(K_n)\right]\right)^2}{\mathbb{E}\left[I_{n,\gamma}(K_n)^2\right]}.$$

Letting n go to infinity in the second inequality at (44) we conclude that

$$\limsup_{n\to\infty}\mathbb{P}\left[I_{n,\gamma}(K_n)=0\right]\le 1-\liminf_{n\to\infty}\frac{\left(\mathbb{E}\left[I_{n,\gamma}(K_n)\right]\right)^2}{\mathbb{E}\left[I_{n,\gamma}(K_n)^2\right]}\le 0,$$

and the zero-law $\lim_{n\to\infty}\mathbb{P}\left[I_{n,\gamma}(K_n)=0\right]=0$ is obtained. □

7.2. A technical lemma

In the proofs of Lemmas 7.1 and 7.2 we repeatedly use the following technical fact.

Lemma 7.3. *For any sequence $m : \mathbb{N}_0 \to \mathbb{N}_0$ with $m_n = \Theta(n)$, we have*

$$\left(1 - \frac{K_n}{m_n}\right)^{m_n} \sim e^{-K_n}. \tag{47}$$

Proof. Recall the elementary decomposition

$$\log(1-x) = -x - \Psi(x) \quad \text{with } \Psi(x) := \int_0^x \frac{t}{1-t}\,dt \tag{48}$$

valid for $0 \le x < 1$. It is easy to check that

$$\lim_{x \downarrow 0} \frac{\Psi(x)}{x^2} = \frac{1}{2}. \tag{49}$$

Under the enforced assumptions we have $m_n = \Theta(n)$ and $K_n = O(\log n)$, whence $\frac{K_n}{m_n} = O(\frac{\log n}{n})$ so that $K_n < m_n$ for all n sufficiently large. On that range, the decomposition (48) yields

$$\left(1 - \frac{K_n}{m_n}\right)^{m_n} = e^{-K_n} \cdot e^{-m_n \Psi\left(\frac{K_n}{m_n}\right)}. \tag{50}$$

Now note that

$$\lim_{n\to\infty} \frac{K_n}{m_n} = 0 \quad \text{and} \quad \lim_{n\to\infty} m_n \left(\frac{K_n}{m_n}\right)^2 = 0.$$

With the help of (49) it is plain that $\lim_{n\to\infty} m_n \Psi\left(\frac{K_n}{m_n}\right) = 0$, and this establishes (47) via (50). □

7.3. A proof of Lemma 7.1

Consider γ in $(0, 1)$. Fix $n = 2, 3, \ldots$ and a positive integer K such that $K < n$. Here as well there is no loss of generality in assuming $n - \lfloor \gamma n \rfloor \ge K$ and $\lfloor \gamma n \rfloor > 1$. Under the enforced assumptions, we get

$$\begin{aligned} \mathbb{E}\left[\chi_{n,\gamma,1}(K)\right] &= \frac{\binom{n-\lfloor \gamma n \rfloor}{K}}{\binom{n-1}{K}} \left(\frac{\binom{n-2}{K}}{\binom{n-1}{K}}\right)^{\lfloor \gamma n \rfloor - 1} \\ &= a(n;K) \cdot \left(1 - \frac{K}{n-1}\right)^{\lfloor \gamma n \rfloor - 1} \end{aligned} \tag{51}$$

with

$$a(n;K) := \frac{(n - \lfloor \gamma n \rfloor)!}{(n - \lfloor \gamma n \rfloor - K)!} \cdot \frac{(n-1-K)!}{(n-1)!}.$$

Now pick a scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$ such that (6) holds for some $c > 0$ and replace K by K_n in (51) according to this scaling. Apply Stirling's formula

$$m! \sim \sqrt{2\pi m} \left(\frac{m}{e}\right)^m \quad (m \to \infty)$$

to the factorials appearing in (51). Doing so readily yields

$$a(n;K_n) \sim \sqrt{\frac{(n - \lfloor \gamma n \rfloor)(n-1-K_n)}{(n - \lfloor \gamma n \rfloor - K_n)(n-1)}} \cdot \alpha_n \beta_n \sim \alpha_n \beta_n \tag{52}$$

under the enforced assumptions on the scaling where we have set

$$\alpha_n := \frac{(n-K_n-1)^{n-K_n-1}}{(n-1)^{n-1}} = \left(1 - \frac{K_n}{n-1}\right)^{n-1} \cdot (n-K_n-1)^{-K_n}$$

and

$$\begin{aligned}\beta_n &:= \frac{(n-\lfloor\gamma n\rfloor)^{n-\lfloor\gamma n\rfloor}}{(n-\lfloor\gamma n\rfloor-K_n)^{n-\lfloor\gamma n\rfloor-K_n}}\\ &= \left(1-\frac{K_n}{n-\lfloor\gamma n\rfloor}\right)^{-(n-\lfloor\gamma n\rfloor)}\cdot(n-\lfloor\gamma n\rfloor-K_n)^{K_n}.\end{aligned}$$

Using (47), first with $m_n = n-1$, then with $m_n = n-\lfloor\gamma n\rfloor$, we obtain

$$\left(1-\frac{K_n}{n-1}\right)^{n-1}\sim e^{-K_n}$$

and

$$\left(1-\frac{K_n}{n-\lfloor\gamma n\rfloor}\right)^{-(n-\lfloor\gamma n\rfloor)}\sim\left(e^{-K_n}\right)^{-1}=e^{K_n},$$

whence

$$\alpha_n\beta_n\sim\left(\frac{n-\lfloor\gamma n\rfloor-K_n}{n-K_n-1}\right)^{K_n}. \tag{53}$$

From (51) and (52) we now conclude that

$$n\mathbb{E}\left[\chi_{n,\gamma,1}(K_n)\right]\sim n\left(1-\frac{K_n}{n-1}\right)^{\lfloor\gamma n\rfloor-1}\cdot\left(\frac{n-\lfloor\gamma n\rfloor-K_n}{n-K_n-1}\right)^{K_n}. \tag{54}$$

A final application of (47), this time with $m_n = n-1$, gives

$$\begin{aligned}\left(1-\frac{K_n}{n-1}\right)^{\lfloor\gamma n\rfloor-1} &= \left(\left(1-\frac{K_n}{n-1}\right)^{n-1}\right)^{\frac{\lfloor\gamma n\rfloor-1}{n-1}}\\ &\sim e^{-\frac{\lfloor\gamma n\rfloor-1}{n-1}K_n}\end{aligned} \tag{55}$$

since $\lim_{n\to\infty}\frac{\lfloor\gamma n\rfloor-1}{n-1}=\gamma>0$. Substituting (55) into (54) we obtain

$$n\mathbb{E}\left[\chi_{n,\gamma,1}(K_n)\right]\sim e^{\zeta_n} \tag{56}$$

with

$$\zeta_n := \log n-\left(\frac{\lfloor\gamma n\rfloor-1}{n-1}-\log\left(\frac{n-\lfloor\gamma n\rfloor-K_n}{n-K_n-1}\right)\right)K_n$$

for all $n = 1, 2, \ldots$. Finally, from the condition (6) on the scaling, we see that

$$\lim_{n\to\infty}\frac{\zeta_n}{\log n}=1-c+c\frac{\log(1-\gamma)}{\gamma}=1-\frac{c}{r(\gamma)}.$$

Thus, $\lim_{n\to\infty}\zeta_n=-\infty$ (resp. ∞) if $c<r(\gamma)$ (resp. $r(\gamma)<c$) and the desired result follows upon using (56). □

7.4. A proof of Lemma 7.2

Fix $n = 3, 4, \ldots$ and a positive integer K with $K < n$. With γ in $(0, 1)$, we again assume that $n-\lfloor\gamma n\rfloor\geq K$ and $\lfloor\gamma n\rfloor>1$. It is a simple matter to check that

$$\mathbb{E}\left[\chi_{n,\gamma,1}(K)\chi_{n,\gamma,2}(K)\right]=\left(\frac{\binom{n-\lfloor\gamma n\rfloor}{K}}{\binom{n-1}{K}}\right)^2\left(\frac{\binom{n-3}{K}}{\binom{n-1}{K}}\right)^{\lfloor\gamma n\rfloor-2}$$

and invoking (51) we readily conclude that

$$\begin{aligned}\frac{\mathbb{E}\left[\chi_{n,\gamma,1}(K)\chi_{n,\gamma,2}(K)\right]}{\left(\mathbb{E}\left[\chi_{n,\gamma,1}(K)\right]\right)^2} &= \left(\frac{\binom{n-3}{K}}{\binom{n-1}{K}}\right)^{\lfloor\gamma n\rfloor-2}\cdot\left(\frac{\binom{n-1}{K}}{\binom{n-2}{K}}\right)^{2(\lfloor\gamma n\rfloor-1)}\\ &= \left(\left(\frac{n-1-K}{n-1}\right)\left(\frac{n-2-K}{n-2}\right)\right)^{\lfloor\gamma n\rfloor-2}\cdot\left(\frac{n-1}{n-1-K}\right)^{2(\lfloor\gamma n\rfloor-1)}\end{aligned}$$

$$
\begin{aligned}
&= \left(\frac{n-2-K}{n-2}\right)^{\lfloor \gamma n \rfloor -2} \cdot \left(\frac{n-1}{n-1-K}\right)^{\lfloor \gamma n \rfloor} \\
&= \left(1-\frac{K}{n-2}\right)^{\lfloor \gamma n \rfloor -2} \cdot \left(1+\frac{K}{n-1-K}\right)^{\lfloor \gamma n \rfloor} \\
&\le e^{-K \cdot E_n(K)}
\end{aligned}
\tag{57}
$$

where we have set

$$
E_n(K) := \frac{\lfloor \gamma n \rfloor - 2}{n-2} - \frac{\lfloor \gamma n \rfloor}{n-1-K}.
$$

Elementary calculations show that

$$
-K \cdot E_n(K) = \frac{\lfloor \gamma n \rfloor}{n-2} \cdot \frac{K(K-1)}{n-1-K} + \frac{2K}{n-2}.
$$

Now pick a scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$ such that (6) holds for some $c > 0$ and replace K by K_n in (57) according to this scaling. It is plain that $\lim_{n\to\infty} K_n E_n(K_n) = 0$, and the conclusion (46) follows. □

8. Proofs of Lemma 4.1, Theorems 4.2 and 4.3

8.1. Basic concentration inequalities for binomial rvs

Fix $n = 2, 3, \ldots$ and a positive integer K with $K < n$. For each $i = 1, \ldots, n$, it follows from (13) that

$$
|\Sigma_{n,i}(K)| = K + B_{n,i}(K) \tag{58}
$$

where $B_{n,i}(K)$ is the rv determined through

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

Under the enforced independence assumptions, the rv $B_{n,i}(K)$ is a binomial rv $\mathrm{Bin}(n-1, \frac{K}{n-1})$ with mean

$$
\mathbb{E}\left[B_{n,i}(K)\right] = (n-1) \cdot \frac{K}{n-1} = K. \tag{59}
$$

Of particular relevance here is the following well-known concentration inequalities for binomial rvs [22, Lemma 1.1, p. 16]: With $H(t) := 1 - t + t \log t$ $(t > 0)$, we have

$$
\mathbb{P}\left[\, B_{n,1}(K) > K + t \,\right] \le e^{-K \cdot H(\frac{K+t}{K})} \tag{60}
$$

and

$$
\mathbb{P}\left[\, B_{n,1}(K) < K - t \,\right] \le e^{-K \cdot H(\frac{K-t}{K})} \tag{61}
$$

where the additional condition $0 < t < K$ is required for (61) to hold. Simple calculations yield

$$
H\left(\frac{K \pm t}{K}\right) = a\left(\pm \frac{t}{K}\right) \tag{62}
$$

on the appropriate ranges as we make use of (16).

Taking the derivative of (16) we find

$$
\frac{d}{d\tau} a(\tau) = \log(1+\tau), \quad \tau > -1.
$$

Therefore, the mapping $\tau \to a(\tau)$ is convex on $(-1, \infty)$, first strictly decreasing on $(-1, 0)$ and then strictly increasing on $(0, \infty)$ with $\lim_{\tau \downarrow -1} a(\tau) = 1$, $a(0) = 0$ and $\lim_{\tau \to \infty} a(\tau) = \infty$, hence $a(\tau) > 0$ on $(-1, 0) \cup (0, \infty)$. Since

$$
\frac{d}{d\tau}\left(a(-\tau) - a(\tau)\right) = -\log(1-\tau^2) > 0, \quad 0 < \tau < 1,
$$

it is easy to check that

$$
a(\tau) < a(-\tau), \quad 0 < \tau < 1. \tag{63}
$$

Please cite this article in press as: O. Yağan, A.M. Makowski, On the scalability of the random pairwise key predistribution scheme: Gradual deployment and key ring sizes, Performance Evaluation (2013), http://dx.doi.org/10.1016/j.peva.2013.03.001

8.2. A proof of Lemma 4.1

Fix $n = 2, 3, \ldots$ and a positive integer K with $K < n$. To take advantage of (60)–(61) we note from (58) that

$$|\Sigma_{n,1}(K)| - 2K = B_{n,1}(K) - K. \tag{64}$$

For each $t > 0$, it then follows that

$$\mathbb{P}\left[\, \left| |\Sigma_{n,1}(K)| - 2K \right| > t \,\right] = \mathbb{P}\left[\, B_{n,1}(K) > K + t \,\right] + \mathbb{P}\left[\, B_{n,1}(K) < K - t \,\right] \tag{65}$$

with $\mathbb{P}\left[B_{n,1}(K) < K - t\right] = 0$ for $t \geq K$. Using (60)–(61) and the definition (17), we readily conclude from (65) that

$$\mathbb{P}\left[\, \left| |\Sigma_{n,1}(K)| - 2K \right| > t \,\right] \leq b\left(\frac{t}{K}\right) e^{-Ka\left(\frac{t}{K}\right)}$$

as we recall the comparison (63). The desired conclusion (18) follows as we replace t by cK in this last inequality. □

8.3. A proof of Theorem 4.2

Fix $n = 2, 3, \ldots$ and a positive integer K with $K < n$. Again using (58) we get

$$\left(\max_{i=1,\ldots,n} |\Sigma_{n,i}(K)|\right) - 2K = \max_{i=1,\ldots,n} \left(B_{n,i}(K) - K\right).$$

As in the proof of Lemma 4.1, for any given $t > 0$, we now find

$$\begin{aligned} \mathbb{P}\left[\, M_n(K) > 2K + t \,\right] &= \mathbb{P}\left[\, \max_{i=1,\ldots,n} \left(B_{n,i}(K) - K\right) > t \,\right] \\ &= \mathbb{P}\left[\, \max_{i=1,\ldots,n} B_{n,i}(K) > K + t \,\right]. \end{aligned} \tag{66}$$

A simple union argument shows that

$$\begin{aligned} \mathbb{P}\left[\, \max_{i=1,\ldots,n} B_{n,i}(K) > K + t \,\right] &= \mathbb{P}\left[\, \cup_{i=1}^{n} [B_{n,i}(K) > K + t] \,\right] \\ &\leq \sum_{i=1}^{n} \mathbb{P}\left[\, B_{n,i}(K) > K + t \,\right] \\ &= n\mathbb{P}\left[\, B_{n,1}(K) > K + t \,\right] \end{aligned} \tag{67}$$

since the rvs $B_{n,1}(K), \ldots, B_{n,n}(K)$ are identically distributed (but not independent). By the first concentration inequality (60) we then conclude that

$$\mathbb{P}\left[\, \max_{i=1,\ldots,n} B_{n,i}(K) > K + t \,\right] \leq e^{\log n - Ka\left(\frac{t}{K}\right)}. \tag{68}$$

Next, consider a scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$ satisfying (21) for some $\lambda > 0$, and select the sequence $t : \mathbb{N}_0 \to \mathbb{R}_+$ given by

$$t_n = cK_n, \quad n = 1, 2, \ldots$$

with $c > 0$. Replacing K and t accordingly by K_n and t_n in (68), we conclude from (66) that

$$\mathbb{P}\left[M_n(K_n) > (2 + c)K_n\right] \leq e^{-(-\log n + K_n a(c))}. \tag{69}$$

Under the enforced assumptions (21) it is easy to check that

$$\begin{aligned} -\log n + K_n a\left(c\right) &= \left(-1 + \lambda a(c) + o(1)\right) \cdot \log n \\ &= \left(h(\lambda; c) + o(1)\right) \cdot \log n \end{aligned}$$

and the bound (22) follows.

Finally pick $\lambda > 0$: Since $a(0) = 0$ and $\lim_{\tau\to\infty} a(\tau) = \infty$, the strict monotonicity of the mapping $\tau \to a(\tau)$ on $\mathbb{R}_+$ implies that the equation

$$\lambda a\left(c\right) = 1, \quad c > 0 \tag{70}$$

has a unique solution, hereafter denoted $c(\lambda)$. This statement is equivalent to $c(\lambda)$ being the only solution to the equation $h(\lambda; c) = 0$, $c > 0$, hence the mapping $c \to h(\lambda; c)$ changes sign only once on $\mathbb{R}_+$. From arguments given earlier, it is clear that

$$1 < \lambda a\left(c\right), \quad c(\lambda) < c.$$

This last statement being equivalent to (24), the proof is now completed. □

Please cite this article in press as: O. Yağan, A.M. Makowski, On the scalability of the random pairwise key predistribution scheme: Gradual deployment and key ring sizes, Performance Evaluation (2013), http://dx.doi.org/10.1016/j.peva.2013.03.001

We conclude by addressing the comments following the statement of Theorem 4.2: By the Implicit Function Theorem, the mapping $\lambda \to c(\lambda)$ is differentiable on $(0, \infty)$. The defining relation (70) for $c(\lambda)$ implies

$$a(c(\lambda)) = \frac{1}{\lambda}, \quad \lambda > 0.$$

Differentiating we find

$$\log(1 + c(\lambda)) \cdot \frac{d}{d\lambda} c(\lambda) = -\frac{1}{\lambda^2}, \quad \lambda > 0$$

and the conclusion $\frac{d}{d\lambda} c(\lambda) < 0$ follows whenever $\lambda > 0$ since then $c(\lambda) > 0$. In short, $\lambda \to c(\lambda)$ is strictly decreasing on $(0, \infty)$—this could also have been established graphically. Also, note that

$$\lambda a(1) = \lambda (2 \log 2 - 1) = \frac{\lambda}{\lambda^\star}, \quad \lambda > 0 \tag{71}$$

where $\lambda^\star$ is given by (25). Thus, $\lambda^\star a(1) = 1$ so that $c(\lambda^\star) = 1$ necessarily, and by the strict monotonicity of $\lambda \to c(\lambda)$ established earlier, it is plain from (71) that $1 < c(\lambda)$ (resp. $c(\lambda) < 1$) if and only if $\lambda < \lambda^\star$ (resp. $\lambda^\star < \lambda$).

8.4. A proof of Theorem 4.3

Fix $n = 2, 3, \ldots$ and a positive integer K with $K < n$, and pick $c > 0$. The arguments that lead to (66)–(67) in the proof of Theorem 4.2 also yield

$$\mathbb{P}[M_n(K) > K + w_n(c)] \le n\mathbb{P}\left[B_{n,1}(K) > w_n(c)\right] \tag{72}$$

where for convenience we have set

$$w_n(c) = c \cdot \frac{\log n}{\log \log n}.$$

Therefore, (28) will be established if we show that

$$n\mathbb{P}\left[B_{n,1}(K) > w_n(c)\right] \le n^{1-c+o(1)}. \tag{73}$$

Note that

$$\lim_{n\to\infty} \frac{w_n(c)}{n} = 0$$

so that $w_n(c) < n$ for all n sufficiently, say $n \ge n^\star$ for some finite integer $n^\star$. On that range, since the rv $B_{n,1}(K)$ is a binomial rv $\mathrm{Bin}(n-1, \frac{K}{n-1})$, we find

$$\begin{aligned}
\mathbb{P}\left[B_{n,1}(K) > w_n(c)\right] &\le \mathbb{P}\left[B_{n,1}(K) \ge \lfloor w_n(c) \rfloor\right] \\
&= \sum_{\ell=\lfloor w_n(c)\rfloor}^{n-1} \binom{n-1}{\ell} \left(\frac{K}{n-1}\right)^\ell \left(1 - \frac{K}{n-1}\right)^{n-1-\ell} \\
&\le \sum_{\ell=\lfloor w_n(c)\rfloor}^{n-1} \binom{n-1}{\ell} \left(\frac{K}{n-1}\right)^\ell \\
&\le \sum_{\ell=\lfloor w_n(c)\rfloor}^{n-1} \left(\frac{e(n-1)}{\ell}\right)^\ell \left(\frac{K}{n-1}\right)^\ell \qquad (74) \\
&= \sum_{\ell=\lfloor w_n(c)\rfloor}^{n-1} \left(\frac{eK}{\ell}\right)^\ell \\
&\le \sum_{\ell=\lfloor w_n(c)\rfloor}^{\infty} \left(\frac{eK}{\lfloor w_n(c)\rfloor}\right)^\ell \qquad (75)
\end{aligned}$$

upon using the bound (35) in (74).

Because $\lim_{n\to\infty} w_n(c) = \infty$, we have $eK < \lfloor w_n(c) \rfloor$ for all n sufficiently large beyond $n^\star$. On that range, the infinite geometric series in (75) is summable, and the inequality

$$\mathbb{P}\left[B_{n,1}(K) > w_n(c)\right] \le \left(1 - \frac{eK}{\lfloor w_n(c)\rfloor}\right)^{-1} \cdot \left(\frac{eK}{\lfloor w_n(c)\rfloor}\right)^{\lfloor w_n(c)\rfloor} \tag{76}$$

follows. From the decomposition (48), we also note that

$$-\log\left(1-\frac{eK}{\lfloor w_n(c)\rfloor}\right)=\frac{eK}{\lfloor w_n(c)\rfloor}+\Psi\left(\frac{eK}{\lfloor w_n(c)\rfloor}\right)\sim\frac{eK}{\lfloor w_n(c)\rfloor}.$$

Therefore,

$$\frac{-\log\left(1-\frac{eK}{\lfloor w_n(c)\rfloor}\right)}{\log n}\sim\frac{eK}{c}\cdot\frac{\log\log n}{(\log n)^2},$$

and we readily conclude that

$$\left(1-\frac{eK}{\lfloor w_n(c)\rfloor}\right)^{-1}=n^{o(1)}. \tag{77}$$

Thus, in view of (75)–(77) we shall be able to conclude (73) if we can show that

$$n\left(\frac{eK}{\lfloor w_n(c)\rfloor}\right)^{\lfloor w_n(c)\rfloor}=n^{1-c+o(1)}. \tag{78}$$

Taking logarithms in the left hand side of (78) we find

$$\begin{aligned}\log\left(n\left(\frac{eK}{\lfloor w_n(c)\rfloor}\right)^{\lfloor w_n(c)\rfloor}\right)&=\log n+\lfloor w_n(c)\rfloor\log eK-\lfloor w_n(c)\rfloor\log\lfloor w_n(c)\rfloor\\&=\left(1+\frac{\lfloor w_n(c)\rfloor}{\log n}\log eK-\frac{\lfloor w_n(c)\rfloor}{\log n}\log\lfloor w_n(c)\rfloor\right)\cdot\log n\\&=(1-c+o(1))\cdot\log n\end{aligned}$$

since

$$\lim_{n\to\infty}\frac{\lfloor w_n(c)\rfloor}{\log n}=0\quad\text{and}\quad\lim_{n\to\infty}\frac{\lfloor w_n(c)\rfloor}{\log n}\cdot\log\lfloor w_n(c)\rfloor=c.$$

This establishes (78), and the proof of Theorem 4.3 is now completed. □

Acknowledgments

This work was supported by NSF Grant CCF-0729093. We thank Dr. H. Chan from CyLab at Carnegie Mellon University for pointing out the Ref. [12] and for several insightful comments regarding the work reported here. We also thank Prof. A. Srinivasan from the Department of Computer Science at the University of Maryland at College Park for making us aware of an earlier version of Theorem 2.1 given in [20]. We also thank the reviewers for their careful reading of the manuscript in its earlier incarnation and for their constructive comments.

References

[1] L. Eschenauer, V.D. Gligor, A key-management scheme for distributed sensor networks, in: Proceedings of the 9th ACM Conference on Computer and Communications Security, CCS 2002, Washington, DC, November 2002, ACM, New York, NY, pp. 41–47.
[2] A. Perrig, J. Stankovic, D. Wagner, Security in wireless sensor networks, Communications of the ACM 47 (2004) 53–57.
[3] D.-M. Sun, B. He, Review of key management mechanisms in wireless sensor networks, Acta Automatica Sinica 12 (2006) 900–905.
[4] Y. Wang, G. Attebury, B. Ramamurthy, A survey of security issues in wireless sensor networks, IEEE Communications Surveys & Tutorials 8 (2006) 2–23.
[5] S.R. Blackburn, S. Gerke, Connectivity of the uniform random intersection graph, Discrete Mathematics 309 (2009) 5130–5140.
[6] R. Di Pietro, L.V. Mancini, A. Mei, A. Panconesi, J. Radhakrishnan, Redoubtable sensor networks, ACM Transactions on Information and System Security 11 (3) (2008) 1–22.
[7] K. Rybarczyk, Diameter, connectivity, and phase transition of the uniform random intersection graph, Discrete Mathematics 311 (2011) 1998–2019.
[8] 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.
[9] O. Yağan, A.M. Makowski, Zero-one laws for connectivity in random key graphs, IEEE Transactions on Information Theory 58 (Number 5) (2012) 2983–2999.
[10] O. Yağan, Performance of the Eschenauer–Gligor key distribution scheme under an ON/OFF channel, IEEE Transactions on Information Theory 58 (Number 6) (2012) 3821–3835.
[11] J. Zhao, O. Yağan, V. Gligor, k-connectivity in secure wireless sensor networks with physical link constraints—the ON/OFF channel model, IEEE Transactions on Information Theory. Available online at: arXiv:1206.1531 [cs.IT] (submitted for publication).
[12] H. Chan, A. Perrig, D. Song, Random key predistribution schemes for sensor networks, in: Proceedings of the 2003 IEEE Symposium on Security and Privacy, SP 2003, Oakland, CA, May 2003, IEEE Computer Society, Washington DC, pp. 197–213.
[13] W. Du, J. Deng, Y.S. Han, P.K. Varshney, A pairwise key pre-distribution scheme for wireless sensor networks, in: Proceedings of the 10th ACM Conference on Computer and Communications Security, CCS 2003, Washington, DC, October 2003, ACM, New York, NY, pp. 42–51.
[14] O. Yağan, A.M. Makowski, Connectivity results for sensor networks under a random pairwise key distribution scheme, in: Proceedings of the IEEE International Symposium on Information Theory, ISIT 2012, Boston, MA, July 2012.

[15] O. Yağan, A.M. Makowski, On the connectivity of sensor networks under random pairwise key predistribution, IEEE Transactions on Information Theory (2013) (in press).

[16] O. Yağan, 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.

[17] O. Yağan, A.M. Makowski, Modeling the pairwise key predistribution scheme in the presence of unreliable links, IEEE Transactions on Information Theory 59 (Number 3) (2013) 1740–1760.

[18] O. Yağan, A.M. Makowski, On the gradual deployment of random pairwise key distribution schemes, in: Proceedings of the 9th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, WiOpt 2011, Princeton, NJ, May 2011.

[19] W. Du, J. Deng, Y.S. Han, P.K. Varshney, J. Katz, A. Khalili, A pairwise key pre-distribution scheme for wireless sensor networks, ACM Transactions on Information and System Security 8 (2) (2005) 228–258.

[20] T.I. Fenner, A.M. Frieze, On the connectivity of random m-orientable graphs and digraphs, Combinatorica 2 (1982) 347–359.

[21] B. Bollobás, Random Graphs, second ed., in: Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, UK, 2001.

[22] M.D. Penrose, Random Geometric Graphs, in: Oxford Studies in Probability, vol. 5, Oxford University Press, New York, NY, 2003.

[23] M. Draief, L. Massoulie, Epidemics and Rumours in Complex Networks, in: London Mathematical Society Lecture Notes Series, vol. 369, Cambridge University Press, Cambridge, UK, 2010.

[24] S. Janson, T. Łuczak, A. Ruciński, Random Graphs, in: Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, USA, 2000.

Osman Yağan received the B.S. degree in Electrical and Electronics Engineering from the Middle East Technical University, Ankara (Turkey) in 2007, and the Ph.D. degree in Electrical and Computer Engineering from the University of Maryland at College Park in 2011.

He was a visiting Postdoctoral Scholar at Arizona State University during Fall 2011. Since December 2011, he has been a Postdoctoral Research Fellow in CyLab at Carnegie Mellon University. His research interests include wireless network security, dynamical processes in complex networks, percolation theory, random graphs and their applications.

Armand M. Makowski received the Licence en Sciences Mathématiques from the Université Libre de Bruxelles in 1975, the M.S. degree in Engineering-Systems Science from U.C.L.A. in 1976 and the Ph.D. degree in Applied Mathematics from the University of Kentucky in 1981. In August 1981, he joined the faculty of the Electrical Engineering Department at the University of Maryland College Park, where he is Professor of Electrical and Computer Engineering. He has held a joint appointment with the Institute for Systems Research since its establishment in 1985.

Armand Makowski was a C.R.B. Fellow of the Belgian-American Educational Foundation (BAEF) for the academic year 1975–76; he is also a 1984 recipient of the NSF Presidential Young Investigator Award and became an IEEE Fellow in 2006.

His research interests lie in applying advanced methods from the theory of stochastic processes to the modeling, design and performance evaluation of engineering systems, with particular emphasis on communication systems and networks.

Please cite this article in press as: O. Yağan, A.M. Makowski, On the scalability of the random pairwise key predistribution scheme: Gradual deployment and key ring sizes, Performance Evaluation (2013), http://dx.doi.org/10.1016/j.peva.2013.03.001