Connectivity in inhomogeneous random key graphs

Osman Yağan
Department of ECE, Carnegie Mellon University
Pittsburgh, PA 15213
Email: oyagan@ece.cmu.edu

***Abstract*—We consider a new random key predistribution scheme for securing *heterogeneous* wireless sensor networks. Each of the n sensors in the network is classified into r classes according to a probability distribution $\mu = \{\mu_1, \ldots, \mu_r\}$. Before deployment, a class-i sensor is assigned K_i cryptographic keys that are selected uniformly at random from a pool of P keys. Once deployed, a pair of sensors can communicate *securely* if and only if they have a key in common. The communication topology of this network is modeled by an *inhomogeneous* random key graph. We establish scaling conditions on the parameters P and $\{K_1, \ldots, K_r\}$ so that this graph is connected with high probability. The result is given in the form of a zero-one law with the number of sensors n growing unboundedly large. Our result is shown to complement and improve those given by Godehardt et al. and Zhao et al. for the same model, therein referred to as the *general random intersection graph*.**

***Index Terms*—Heterogeneous wireless sensor networks; key predistribution; random graphs; connectivity.**

I. Introduction

Random key graphs have been introduced by Yağan and Makowski [22] to study the Eschenauer-Gligor (EG) random key predistribution scheme [8], a widely recognized solution for securing wireless sensor network (WSN) communications [4], [7]. Denoted by $\mathbb{G}(n, K, P)$, a random key graph is constructed on the vertices $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ as follows. Each vertex v_i is assigned *independently* a set Σ_i of K cryptographic keys that are selected uniformly at random from a pool of size P. Any pair of vertices v_i, v_j are then deemed *adjacent* if they share a key, i.e., if $\Sigma_i \cap \Sigma_j \neq \emptyset$. Random key graphs have recently received attention in a wide range of areas including modeling small world networks [21], recommender systems [11], and clustering and classification analysis [10]; they are also referred to as *uniform random intersection graphs* in the literature. Properties that have been studied include absence of isolated nodes [22], connectivity [13], [22], [24], k-connectivity [26], and k-robustness [25], among others.

This paper is the second in a series of publications (the first one is [18]) where we introduce a variation of the EG scheme that is more suitable for *heterogeneous* WSNs. Our motivation is that many military and commercial WSN applications will consist of heterogeneous nodes [14], [16] with varying level of resources (e.g., computational, memory, power) and possibly with varying level of security and connectivity requirements. As a result of this heterogeneity, it may no longer be sensible to assign the same number of keys to all sensors in the network as prescribed by the EG scheme. Instead, we consider a scheme where the number of keys assigned to each sensor is drawn independently from $\boldsymbol{K} = \{K_1, \ldots, K_r\}$ according to a probability distribution $\boldsymbol{\mu} = \{\mu_1, \ldots, \mu_r\}$. Put differently, each vertex v_x is independently assigned to a priority class-i with probability $\mu_i > 0$ and then receives a key ring with the size K_i associated with this class. As before, we assume that once its size is fixed, the key ring Σ_x is constructed by sampling the key pool randomly and without replacement.

Let $\mathbb{G}(n; \boldsymbol{\mu}, \boldsymbol{K}, P)$ denote the random graph induced by the heterogeneous key predistribution scheme described above, where again a pair of nodes are adjacent as long as they share a key; see Section II for precise definitions. Inspired by the recently studied inhomogeneous Erdős-Rényi (ER) graphs [3], [5], we refer to this graph as the *inhomogeneous* random key graph. This model was first introduced by the author in [18], where zero-one laws for the property of absence of isolated nodes were presented. The main goal of the current paper is to extend these results to the *connectivity* of $\mathbb{G}(n; \boldsymbol{\mu}, \boldsymbol{K}, P)$. Namely, we seek to understand how the parameters $n, \boldsymbol{\mu}, \boldsymbol{K}, P$ should behave so that the resulting graph is connected almost surely. Such results can be useful in deriving guidelines for designing heterogenous WSNs so that they are securely connected. By comparison with the results for the standard random key graph, they can also shed light on the effect of heterogeneity on the connectivity properties of WSNs.

Our main result is a zero-one law for the connectivity in $\mathbb{G}(n; \boldsymbol{\mu}, \boldsymbol{K}, P)$ (see Theorem 1). Namely, we scale the parameters $\boldsymbol{K}$ and P and provide critical conditions on the scaling such that the resulting graph is almost surely connected and almost surely *not* connected, respectively, when the number of nodes n goes to infinity. The critical scaling is shown to coincide that obtained for the absence of isolated nodes in $\mathbb{G}(n; \boldsymbol{\mu}, \boldsymbol{K}, P)$ [18], meaning that absence of isolated nodes and connectivity are asymptotically equivalent properties for the inhomogeneous random key graph. Other well-known models that exhibit the same behavior include ER graphs [2], random key graphs [22], and random geometric graphs [12].

Our result is compared with those obtained by Zhao et al. [25] and Godehardt et al. [10] for the k-connectivity and connectivity, respectively, of $\mathbb{G}(n; \boldsymbol{\mu}, \boldsymbol{K}, P)$; there $\mathbb{G}(n; \boldsymbol{\mu}, \boldsymbol{K}, P)$ was referred to as a *general* random intersection graph. We show that earlier results are constrained to parameter ranges that are unlikely to be feasible in real world WSN implementations due to excessive memory requirement or very limited resiliency against adversarial attacks. On the contrary, our results cover parameter ranges that are widely regarded as feasible for most WSNs; see Section III-C for details.

A rather surprising conclusion derived from our main result is that the minimum key ring size in the network has a significant impact on the connectivity of $\mathbb{G}(n;\boldsymbol{\mu},\boldsymbol{K},P)$. In particular, for the (homogeneous) random key graph $\mathbb{G}(n;K,P)$ the critical threshold for connectivity is known [13], [22] to be given by $\frac{K^2}{P} \sim c\frac{\log n}{n}$ and the resulting graph is asymptotically almost surely connected (resp. not connected) if $c>1$ (resp. $c<1$). For the inhomogeneous random key graph $\mathbb{G}(n;\boldsymbol{\mu},\boldsymbol{K},P)$ one would be tempted to think that an equivalent result holds under the scaling $\frac{K_{\text{avg}}^2}{P} \sim c\frac{\log n}{n}$, with $K_{\text{avg}}=\sum_{j=1}^r \mu_j K_j$ denoting the mean key ring size. Instead, we show that the zero-one law for connectivity holds under $\frac{K_{\min}K_{\text{avg}}}{P} \sim c\frac{\log n}{n}$, where $K_{\min}$ stands for the minimum of $\{K_1,\ldots,K_r\}$; see Corollary 2. This implies that under the heterogeneous scheme, the mean number of keys required per sensor node to achieve connectivity can be significantly larger than that required in the homogeneous case. For instance, the expense of allowing an arbitrarily small fraction of sensors to keep half as many keys as in the homogeneous case would be to increase the average key ring size by two-fold.

We close with a word on notation. All limiting statements are understood with the number of sensor nodes n going to infinity. An event is said to hold *with high probability* (whp) if it holds with probability 1 as $n\to\infty$. With arbitrary sequences $\{a_n\},\{b_n\}$, we use $a_n=o(b_n)$, $a_n=w(b_n)$, $a_n=O(b_n)$, $a_n=\Omega(b_n)$, and $a_n=\Theta(b_n)$, with their meaning in the standard Landau notation. We also use $a_n\sim b_n$ to denote the asymptotic equivalence $\lim_{n\to\infty} a_n/b_n=1$.

II. Model Definitions

Consider a network that consists of n sensor nodes labeled as $v_1,\ldots,v_n$. The main idea is to classify the nodes into r sets (e.g., depending on their level of importance) and then ti assign different number of cryptographic keys to sensors based on their class. Assume that each of the n nodes in the network are independently assigned to a class according to some probability distribution $\boldsymbol{\mu}:\{1,\ldots,r\}\to(0,1)$. Namely, with t_x denoting the class (or, type) of node v_x, we have

$$\mathbb{P}\left[t_\ell=i\right]=\mu_i>0, \qquad i=1,\ldots,r,$$

for each $\ell=1,\ldots,n$. Then, a class-i node is assigned K_i keys that are selected uniformly at random from a pool of size P, for each $i=1,\ldots,r$. It is further assumed that the rvs $\Sigma_1,\ldots,\Sigma_n$ are independent and identically distributed.

Let $\boldsymbol{K}=(K_1,\ldots,K_r)$ and $\boldsymbol{\mu}=(\mu_1,\ldots,\mu_r)$. Without loss of generality we assume that $K_1\le K_2\le\cdots\le K_r$. Consider a random graph $\mathbb{G}$ defined on the vertex set $\mathcal{V}=\{v_1,\ldots,v_n\}$ such that two nodes v_x and v_y are adjacent, denoted $v_x\sim v_y$, if they have at least one key in common in their corresponding key rings. Namely, we have

$$v_x\sim v_y \qquad \text{if} \qquad \Sigma_x\cap\Sigma_y\neq\emptyset. \tag{1}$$

The adjacency condition (1) defines the *inhomogeneous random key graph*, hereafter denoted $\mathbb{G}(n;\boldsymbol{\mu},\boldsymbol{K},P)$. The name is reminiscent of the recently studied inhomogeneous random graph [3] model where nodes are again divided into r classes, and a class i node and a class j node are connected with probability p_{ij}, independent of everything else. This independence disappears in the inhomogeneous random key graph case, but one can still compute p_{ij} as

$$p_{ij}:=1-\frac{\binom{P-K_i}{K_j}}{\binom{P}{K_j}}, \quad i,j=1,\ldots,r. \tag{2}$$

In view of (2), our key predistribution scheme results in higher priority nodes (i.e., nodes with more assigned keys) connecting with each other with higher probability; see [19]. In presenting our results below, we shall make use of the *mean* probability of edge occurrence for each node class. Namely, we define

$$\lambda_i:=\sum_{j=1}^r p_{ij}\mu_j, \quad i=1,\ldots,r. \tag{3}$$

It is easy to see that the mean number of edges incident on a node (i.e., the *degree* of a node) of class-i is given by $(n-1)\lambda_i$.

Throughout, we assume that the number of classes r is fixed and do not scale with n, and so are the probabilities $\mu_1,\ldots,\mu_r>0$. All other parameters are scaled with n, and we are interested in the properties of the resulting inhomogeneous random key graph as n grows unboundedly large.

III. Main Results and Discussion

A. The results

To fix the terminology, we refer to any mapping $K_1,\ldots,K_r,P:\mathbb{N}_0\to\mathbb{N}_0^{r+1}$ as a *scaling* as long as

$$1\le K_{1,n}\le K_{2,n}\le\cdots\le K_{r,n}<P_n \tag{4}$$

holds for $n=2,3,\ldots$. Let $\boldsymbol{K}_n=(K_{1,n},K_{2,n},\ldots,K_{r,n})$. Our main result, presented below, is a zero-one law for the connectivity of inhomogeneous random key graphs.

Theorem 1. *Consider a probability distribution $\boldsymbol{\mu}=(\mu_1,\ldots,\mu_r)$ with $\mu_i>0$ for $i=1,\ldots,r$, and a scaling $K_1,\ldots,K_r,P:\mathbb{N}_0\to\mathbb{N}_0^{r+1}$ such that*

$$\lambda_1(n)\sim c\frac{\log n}{n} \tag{5}$$

for some $c>0$. Under the assumptions

$$P_n=\Omega(n) \tag{6}$$

and

$$\frac{(K_{1,n})^2}{P_n}=w\left(\frac{1}{n}\right), \tag{7}$$

we have

$$\lim_{n\to\infty}\mathbb{P}\left[\mathbb{G}(n;\boldsymbol{\mu},\boldsymbol{K}_n,P_n)\text{ is connected}\right]=\begin{cases}0 & \text{if } c<1\\ 1 & \text{if } c>1.\end{cases} \tag{8}$$

The proof of Theorem 1 is omitted here due to space limitations. All details can be found in [19].

In words, Theorem 1 states that the inhomogeneous random key graph $\mathbb{G}(n;\boldsymbol{\mu},\boldsymbol{K}_n,P_n)$ is connected whp if the mean degree of "the nodes that have the least number of keys" is scaled as $(1+\epsilon)\log n$ for some $\epsilon>0$; in view of [19, Proposition 4.1], the nodes that are assigned the least number of keys have the *minimum* mean-degree in the graph. On the other hand, if this minimal mean degree scales like $(1-\epsilon)\log n$ for some $\epsilon>0$, then whp $\mathbb{G}(n;\boldsymbol{\mu},\boldsymbol{K}_n,P_n)$ is *not* connected. The additional conditions (6) and (7) are enforced here merely for technical reasons and are required only for the one-law part of the connectivity result, Theorem 1. A detailed discussion on these additional conditions is given in Section III-B, where we explain why they are likely to hold in many real-world WSN applications. There, we also discuss how and when these conditions can be relaxed or replaced by milder conditions.

In [18, Theorem 1], the author established an analog of Theorem 1 for the "absence of isolated nodes" property in inhomogeneous random key graphs. Namely, they showed that under the scaling (5), $\mathbb{G}(n;\boldsymbol{\mu},\boldsymbol{K}_n,P_n)$ has no isolated nodes (resp. has at least one isolated node) whp if $c>1$ (resp. if $c<1$); the conditions (6) and (7) were not needed for this result to hold. With this in mind, Theorem 1 complements and extends the absence of isolated nodes result given in [18] to the stronger (and more desired) property of connectivity. The results given here also demonstrate that the inhomogeneous random key graph provides one more example random graph model where the properties of absence of isolated nodes and connectivity are asymptotically equivalent. Other well-known examples include Erdős-Rényi graphs [2], random key graphs [22], random geometric graphs [12], intersection of random key graphs and ER graphs [17], and intersection of random K-out graphs and ER graphs [15], [23].

Our result is also analogous to the recent findings by Levroye and Freiman [5] for the connectivity of inhomogeneous Erdős-Rényi graph model, where nodes are classified into r classes independently according to a probability distribution $\boldsymbol{\mu}$ and an edge is drawn between a class-i and a class-j node with probability $p_{ij}(n)$ independent of everything else. With $\lambda_i(n)$ defined as $\lambda_i(n):=\sum_{j=1}^r p_{ij}(n)\mu_j$, their result states that if $\min_{i=1,\ldots,r}\lambda_i(n)\sim c\log n/n$ then with $c>1$ (resp. $c<1$) the corresponding graph is connected (resp. not connected) whp, under some additional technical conditions.

We now present a corollary of Theorem 1 under a different scaling condition than (5). This alternative formulation makes it easier to derive design guidelines for *dimensioning* heterogeneous key predistribution schemes, namely in adjusting key ring sizes $K_1,\ldots,K_r$ and probabilities $\mu_1,\ldots,\mu_r$ such that the resulting network is connected whp.

Corollary 2. *Consider a probability distribution $\boldsymbol{\mu}=(\mu_1,\ldots,\mu_r)$ with $\mu_i>0$ for $i=1,\ldots,r$ and a scaling $K_1,\ldots,K_r,P:\mathbb{N}_0\to\mathbb{N}_0^{r+1}$. Let $|\Sigma|_n$ denote a rv that takes the value $K_{i,n}$ with probability μ_i for each $i=1,\ldots,r$. If*

$$\frac{K_{1,n}\mathbb{E}\left[|\Sigma|_n\right]}{P_n}\sim c\frac{\log n}{n} \tag{9}$$

holds for some $c>0$, then we have the zero-one law (8) if the additional conditions (6) and (7) are also satisfied.

A proof of Corollary 2 is given in [19], where we show that the scaling conditions (5) and (9) are indeed *equivalent* to each other, meaning that one can obtain Theorem 1 from Corollary 2, and vice versa. We remark that $\mathbb{E}\left[|\Sigma|_n\right]$ gives the mean number of keys assigned to a sensor in the network. With this in mind, Corollary 2 provides various design choices to ensure that resulting network is connected. One just has to set the minimum and average key ring sizes such that their multiplication scales as $(1+\epsilon)\frac{P_n\log n}{n}$ for some $\epsilon>0$, and has to ensure that the additional conditions (6)-(7) are satisfied.

To compare with the homogeneous random key predistribution scheme, set $r=1$ and consider a universal key ring size K_n in Corollary 2. This leads to zero-one laws for connectivity in the standard random key graph $\mathbb{G}(n;K_n,P_n)$. Namely, with

$$\frac{K_n^2}{P_n}\sim c\frac{\log n}{n},\qquad c>0 \tag{10}$$

analogs of (8) are obtained for $\mathbb{G}(n;K_n,P_n)$; these results had already been established in [22] in stronger forms. An interesting observation is that minimum key ring size has a dramatic impact on the connectivity properties of inhomogeneous random key graph. To provide a simple and concrete example, set $P_n=n\log n$. In the homogeneous case, we see from (10) that the universal key ring size has to scale as $K_n=(1+\epsilon)\log n$ for some $\epsilon>0$ to ensure that the network is connected. In the heterogeneous case, one gains the flexibility of having a positive fraction of sensors in the network with substantially smaller number of keys. However, from Corollary 2 we see that this comes at the expense of having to assign a substantially larger key rings to a positive fraction of other sensors. To give a concrete example, we see from (11) that the minimum key ring size $K_{1,n}$ can be kept on the order of $O(\sqrt{\log n})$ and connectivity can still be achieved if the mean key ring size is $O((\log n)^{1.5})$.

B. Comments on the technical conditions (6)-(7)

We now provide a detailed discussion on the technical conditions (6) and (7) enforced in Theorem 1. We will focus on i) the feasibility of these additional conditions for real-world WSN implementations, and ii) when and how they can be replaced with milder conditions.

We start with the condition (6) that states the key pool size grows at least linearly with the network size n. In terms of applicability in the context of heterogeneous key predistribution schemes in WSNs, this condition is not stringent at all. In fact, it is often needed that key pool size P_n be much larger than the network size n [6], [8] as otherwise the network will be extremely vulnerable against node capture attacks. From a technical point of view, the case where $P_n=\Omega(n)$ is also the more interesting and challenging one as compared to the case where $P_n=o(n)$. For instance, when $P_n=O(n^\delta)$ for some $0<\delta<1/2$, the inhomogeneous random key graph $\mathbb{G}(n;\boldsymbol{\mu},\boldsymbol{K}_n,P_n)$ can be shown to be connected for any $\boldsymbol{\mu}$ as long as $K_{1,n}\geq 2$; see [20, Lemma 8.1] for a proof of a similar

result for the standard random key graph. This means that if $P_n = O(n^\delta)$ with $\delta < 1/2$, even two keys per sensor node is enough to get network connectivity whp. Finally, we remark that the scaling condition (5) or its equivalent (9) already implies that $P_n = \Omega(\frac{n}{\log n})$ since $K_{1,n}\mathbb{E}\left[|\Sigma|_n\right] \geq 1$.

Next, we look at the condition (7) and start with discussing possible relaxations. First of all, (7) is stronger than what is actually needed for our proof to work; it is enforced to enable a shorter proof and an easier exposition of the main result. As discussed in [19], we can replace (7) with

$$\begin{cases} \frac{K_{1,n}^2}{P_n} \geq \frac{2\log 2 + \log(1-\mu_r)+\epsilon}{\beta\nu} \frac{1}{n} \text{ and } K_{1,n} = w(1), & \text{if } \mu_r \leq 0.75 \\ \frac{K_{1,n}^2}{P_n} = \Omega\left(\frac{1}{n(\log n)^M}\right) \text{ and } K_{1,n} = w(1), & \text{if } \mu_r > 0.75 \end{cases} \tag{11}$$

for any $\epsilon > 0$ and any finite integer M; here $\beta > 0$ and $\nu > 0$ are variables specified in the proof of Theorem 1.

As we look at (11), we see that $K_{1,n} = w(1)$ is needed for any μ_r. In fact, this condition can easily be satisfied in real-world WSN implementations given that key ring sizes on order of $O(\log n)$ are regarded as feasible for most sensor networks [6]. Considered in combination with (9), other conditions enforced in (11) bound the *variability* in the key ring sizes used in the network. In particular, given that

$$\frac{\mathbb{E}\left[|\Sigma|_n\right]}{K_{1,n}} = \frac{\frac{K_{1,n}\mathbb{E}[|\Sigma|_n]}{P_n}}{\frac{(K_{1,n})^2}{P_n}} = \Theta\left(\frac{\log n}{n}\right)\left(\frac{(K_{1,n})^2}{P_n}\right)^{-1},$$

(11) implies $\frac{\mathbb{E}[|\Sigma|_n]}{K_{1,n}} = O\left(\log n\right)$ when $\mu_r \leq 0.75$ and $\frac{\mathbb{E}[|\Sigma|_n]}{K_{1,n}} = O\left((\log n)^M\right)$ when $\mu_r > 0.75$. Thus, we see that when more than 75 % of the sensors receive the largest key rings, one can afford to use much smaller key rings for the remaining sensors, as compared to the case when $\mu_r \leq 0.75$.

Collecting, while conditions enforced in (11) take away from the flexibility of assigning *very small* key rings to a certain fraction of sensors (as we were allowed to do for the absence of isolated nodes [18], [19]), they can still be satisfied easily in most real-world implementations. To provide a concrete example, one can set $P_n = n\log n$ and have $K_{1,n} = (\log n)^{1/2+\epsilon}$ and $\mathbb{E}\left[|\Sigma|_n\right] = (1+\epsilon)(\log n)^{3/2-\epsilon}$ with any $\epsilon > 0$; in view of Theorem 1 and (11) the resulting network will be connected whp. With the same P_n, it is possible to have much smaller $K_{1,n}$ when $\mu_r > 0.75$. For example, we can have $K_{1,n} = \log\log\cdots\log n$ and $\mathbb{E}\left[|\Sigma|_n\right] = \Omega((\log n)^2)$. Of course, one can also have all key ring sizes on the same order and set $K_{1,n} = c_1 \log n$ and $\mathbb{E}\left[|\Sigma|_n\right] = c_2 \log n$ with $c_1 c_2 > 1$, to obtain a connected WSN whp.

C. Comparison with related work

The model $\mathbb{G}(n; \boldsymbol{\mu}, \boldsymbol{K}_n, P_n)$ considered here is also known as *general random intersection graph* in the literature; e.g., see [1], [9], [25]. To the best of our knowledge this model has been first considered by Godehardt and Jaworski [9] and by Goderhardt et al. [10]. Results for both the absence of isolated nodes and graph connectivity have been established; see below for a comparison of these results with ours. Recently, Zhao et al. [25] established results for the k-connectivity and k-robustness of the general random intersection graph.

We now compare our results with those established in the literature. Our main argument is that previous results for the connectivity of inhomogeneous random key graphs are constrained to very narrow parameter ranges that are impractical for wireless sensor network applications. In particular, we will argue below that the result by Zhao et al. [25] is restricted to *very large* key ring sizes, rendering them impractical for resource-constrained sensor networks. On the other hand, the results by Godehardt et al. [1], [9] focus on fixed key ring sizes that do not grow with the network size n. As a consequence, in order to ensure connectivity, their result requires a key pool size P_n that is *much smaller* than typically prescribed for security and resiliency purposes.

To fix the terminology, let $\mathcal{D}_n : \{1, 2, \ldots, P_n\} \to [0, 1]$ be the probability distribution used for drawing the *size* of the key rings $\Sigma_1, \ldots, \Sigma_n$; as before, once its size is fixed a key ring is formed by sampling a pool of size P_n randomly without replacement. The graph $\mathbb{G}(n; \mathcal{D}_n, P_n)$ is defined on the vertices $\{v_1, \ldots, v_n\}$ and contains an edge between any pair of nodes v_x and v_y as long as $\Sigma_x \cap \Sigma_y \neq \emptyset$. The model $\mathbb{G}(n; \boldsymbol{\mu}, \boldsymbol{K}_n, P_n)$ considered here constitutes a special case of $\mathbb{G}(n; \mathcal{D}_n, P_n)$ under the assumption that the support of $\mathcal{D}_n$ has a fixed size of r. With these in mind, we now state the results by Zhao et al. [25] and Goderhardt et al. [10], consecutively.

Theorem 3. *[25, Theorem 1] Consider a general random intersection graph $\mathbb{G}(n, \mathcal{D}_n, P_n)$. Let $|\Sigma|_n$ be a random variable following the distribution $\mathcal{D}_n$. With α_n defined via*

$$\frac{\mathbb{E}[|\Sigma|_n]^2}{P_n} = \frac{\log n + (k-1)\log\log n + \alpha_n}{n}, \tag{12}$$

if $P_n = \Omega(n)$, $var[|\Sigma|_n] = o\left(\frac{\mathbb{E}[|\Sigma|_n]^2}{n(\log n)^2}\right)$, and $\alpha_n = o(\log n)$,

$$\lim_{n\to\infty} \mathbb{P}\left[\mathbb{G}(n, \mathcal{D}_n, P_n) \text{ is } k\text{-connected}\right] = \begin{cases} 0 & \text{if } \alpha_n \to -\infty \\ 1 & \text{if } \alpha_n \to \infty. \end{cases}$$

Theorem 4. *[10, Theorem 2] Consider a general random intersection graph $\mathbb{G}(n, \mathcal{D}, P_n)$, where $\mathcal{D}(\ell) = 0$ for all $\ell > L$ and $\ell = 0$. Namely, all key ring sizes are bound to be on the interval $[1, L]$. Let $|\Sigma|$ be a random variable following the probability distribution $\mathcal{D}$. Then if*

$$\frac{n}{P_n}(\mathbb{E}\left[|\Sigma|\right] - \mathcal{D}(1)) - \log P_n \to \infty \tag{13}$$

then $\lim_{n\to\infty} \mathbb{P}\left[\mathbb{G}(n, \mathcal{D}, P_n) \text{ is connected}\right] = 1$.

We now argue why the results established in Theorems 3 and 4 are *not* likely to be applicable for real-world sensor networks. First, Theorem 4 focuses on the case where all possible key rings have a finite size that do not scale with n. With $\mathbb{E}\left[|\Sigma|\right]$ fixed, the scaling condition (13) clearly requires

$$P_n = O\left(n/\log n\right). \tag{14}$$

In contrast with (14), it is often needed that key pool size P_n be much larger than the network size n [6], [8] as otherwise

the network will be extremely vulnerable against node capture attacks. In fact, one can see that with (14) in effect, an adversary can compromise a significant portion of the key pool (and, hence network communication) by capturing $o(n)$ nodes.

We now focus on Theorem 3, where the major problem arises from the assumption

$$var[|\Sigma|_n] = o\left(\frac{\mathbb{E}[|\Sigma|_n]^2}{n(\log n)^2}\right). \tag{15}$$

For the model to be deemed as *inhomogeneous* random key graph, the variance of the key ring size should be non-zero. Given that key ring sizes are integer-valued, even the simplest case of assigning either K or $K+1$ keys to each node with some probabilities μ and $1-\mu$, respectively, will lead to $var[|\Sigma|] = \mu(1-\mu) > 0$ (with $0 < \mu < 1$). Therefore, (15) can only be satisfied if $\frac{\mathbb{E}[|\Sigma|_n]^2}{n(\log n)^2} = w(1)$, or, equivalently

$$\mathbb{E}\left[|\Sigma|_n\right] = w\left(\sqrt{n}\log n\right). \tag{16}$$

Put differently, Theorem 3 enforces *mean* key ring size to be much larger than $\sqrt{n}\log n$. However, a typical wireless sensor network will consist of a very large number of sensors, each with very limited memory and computational capability [6], [8]. As a result, key rings with size $w(\sqrt{n}\log n)$ are unlikely to be implementable in most practical network deployments. In fact, it was suggested by Di Pietro et al. [6] that key rings with size $O(\log n)$ are acceptable for sensor networks.

In comparison, our result Theorem 1 does not require either of the unrealistic conditions (14) or (16). To see this, note that the scaling condition (5) implies (see [19, Lemma 4.3])

$$\frac{K_{1,n}K_{r,n}}{P_n} = \Theta\left(\frac{\log n}{n}\right). \tag{17}$$

Obviously, this condition does not require (14), and in fact already enforces $P_n = \Omega(n/\log n)$. Also, the additional conditions (6)-(7) of our connectivity result and (17) can be satisfied simultaneously without requiring the prohibitively large key ring sizes given at (16). To provide concrete examples, we can use $P_n = \Theta(n\log n)$, $K_{1,n} = \Theta(\log n)$ and $K_{r,n} = \Theta(\log n)$, or $P_n = \Theta(n\log n)$, $K_{1,n} = \Theta(\sqrt{\log n})$ and $K_{r,n} = \Theta((\log n)^{3/2})$. With proper choice of constants in these scalings, we will ensure that i) the resulting WSN is connected whp; ii) the key pool size is much larger than the network so that the resulting WSN has good level of resiliency against node capture attacks; and iii) the maximum key ring size used in the network is on the order of the ranges $\log n$ or $(\log n)^{3/2}$ that are usually regarded as feasible [6], [8]; these choices also lead to a much smaller mean key ring size than that prescribed in (16).

ACKNOWLEDGEMENTS

This work has been supported in part by the Department of Electrical and Computer Engineering at Carnegie Mellon University (CMU), by a Berkman Faculty Development Grant from CMU, and a generous gift from Persistent Systems, Inc.

REFERENCES

[1] M. Bloznelis, J. Jaworski, and K. Rybarczyk. Component evolution in a secure wireless sensor network. *Netw.*, 53:19–26, January 2009.

[2] B. Bollobás. *Random graphs*, volume 73. Cambridge university press, 2001.

[3] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures and Algorithms*, 33(1):3–122, 2007.

[4] H. Chan, A. Perrig, and D. Song. Random key predistribution schemes for sensor networks. In *Proc. IEEE Symposium on Security and Privacy*, May 2003.

[5] L. Devroye and N. Fraiman. Connectivity of inhomogeneous random graphs. *Random Structures & Algorithms*, 45(3):408–420, 2014.

[6] R. Di Pietro, L. V. Mancini, A. Mei, A. Panconesi, and J. Radhakrishnan. Redoubtable sensor networks. *ACM Trans. Inf. Syst. Secur.*, 11(3):13:1–13:22, 2008.

[7] W. Du, J. Deng, Y. Han, S. Chen, and P. Varshney. A key management scheme for wireless sensor networks using deployment knowledge. In *Proc. INFOCOM*, 2004.

[8] L. Eschenauer and V. Gligor. A key-management scheme for distributed sensor networks. In *Proc. ACM CCS*, 2002.

[9] E. Godehardt and J. Jaworski. Two models of random intersection graphs for classification. In *Exploratory Data Analysis in Empirical Research*, pages 67–81. Springer Berlin Heidelberg, 2003.

[10] E. Godehardt, J. Jaworski, and K. Rybarczyk. Random intersection graphs and classification. In *Advances in Data Analysis*, pages 67–74. Springer Berlin Heidelberg, 2007.

[11] P. Marbach. A lower-bound on the number of rankings required in recommender systems using collaborativ filtering. In *Proc. IEEE CISS*, 2008.

[12] M. Penrose. *Random Geometric Graphs*. Oxford University Press, 2003.

[13] K. Rybarczyk. Diameter, connectivity and phase transition of the uniform random intersection graph. *Discrete Mathematics*, 311, 2011.

[14] C.-H. Wu and Y.-C. Chung. Heterogeneous wireless sensor network deployment and topology control based on irregular sensor model. In *Advances in Grid and Pervasive Computing*, volume 4459, pages 78–88. Springer Berlin Heidelberg, 2007.

[15] O. Yağan and A. M. Makowski. Designing securely connected wireless sensor networks in the presence of unreliable links. In *IEEE International Conference on Communications (ICC 2011)*, pages 1–5, 2011.

[16] M. Yarvis, N. Kushalnagar, H. Singh, A. Rangarajan, Y. Liu, and S. Singh. Exploiting heterogeneity in sensor networks. In *Proceedings IEEE INFOCOM 2005*, volume 2, pages 878–890 vol. 2, March 2005.

[17] O. Yağan. Performance of the Eschenauer-Gligor key distribution scheme under an on/off channel. *IEEE Transactions on Information Theory*, 58(6):3821–3835, 2012.

[18] O. Yağan. Absence of isolated nodes in inhomogeneous random key graphs. In *Proc. Allerton Conference on Communication, Control, and Computing*, October 2015.

[19] O. Yağan. Zero-one laws for connectivity in inhomogeneous random key graphs. *arXiv preprint arXiv:1508.02407*, 2015.

[20] O. Yağan and A. Makowski. Connectivity in random graphs induced by a key predistribution scheme - small key pools. In *Information Sciences and Systems (CISS), Annual Conference on*, pages 1–6, March 2010.

[21] O. Yağan and A. M. Makowski. Random key graphs – can they be small worlds? In *Proc. International Conference on Networks and Communications (NETCOM)*, pages 313 –318, December 2009.

[22] O. Yağan and A. M. Makowski. Zero–one laws for connectivity in random key graphs. *IEEE Transactions on Information Theory*, 58(5):2983–2999, May 2012.

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

[24] J. Zhao, O. Yagan, and V. Gligor. Connectivity in secure wireless sensor networks under transmission constraints. In *52nd Annual Allerton Conference on Communication, Control, and Computing*, 2014.

[25] J. Zhao, O. Yağan, and V. Gligor. On the strengths of connectivity and robustness in general random intersection graphs. In *IEEE Annual Conference on Decision and Control*, pages 3661–3668, Dec 2014.

[26] J. Zhao, O. Yağan, and V. Gligor. k-connectivity in random key graphs with unreliable links. *IEEE Transactions on Information Theory*, 61(7):3810–3836, July 2015.