THE INSTITUTE FOR SYSTEMS RESEARCH

ISR TECHNICAL REPORT 2010-10

On random graphs associated with a pairwise key distribution scheme

Osman Yagan, Armand Makowski

On random graphs associated with a pairwise key distribution scheme

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

***Abstract*— The pairwise key distribution scheme of Chan et al. was proposed as an alternative to the key distribution scheme of Eschenauer and Gligor to enable network security in wireless sensor networks. We consider the random graph induced by this pairwise scheme under the assumption of full visibility, and show the existence of a zero-one law for graph connectivity.**

Keywords: Wireless sensor networks, Security, Key predistribution, Random key graphs, Connectivity, Zero-one laws.

I. Introduction

Wireless sensor networks (WSNs) are distributed collections of sensors with limited capabilities for computations and wireless communications. Security is expected to be a key challenge for WSNs deployed in hostile environments where communications are monitored, and nodes are subject to capture and surreptitious use by an adversary. However, traditional key exchange and distribution protocols have been found inadequate for large-scale WSNs, e.g., see [6], [11], [13] for discussions of some of the challenges.

Recently *random* key predistribution schemes have been proposed to address some of the difficulties. The idea of randomly assigning secure keys to the sensor nodes prior to network deployment was first introduced by Eschenauer and Gligor [6]. Under full visibility, i.e., when nodes are all within communication range of each other, the EG scheme induces a random graph referred to as a *random key graph* (also known in the literature as a *uniform* random intersection graph [1]). The conditions on the graph parameters to ensure the absence of isolated nodes have been obtained independently in [1], [14], while the references [1], [4], [12], [15], [16] discuss zero-one laws for connectivity in random key graphs. Needless to say, the full visibility assumption does away with the wireless nature of the communication infrastructure supporting WSNs. In return, this simplification makes it possible to focus on how randomizing the key selections affects the establishment of a secure network, and the connectivity results for the underlying random key graph then provide helpful guidelines to dimension the EG scheme.

Following the original work of Eschenauer and Gligor, a number of other key distribution schemes have been suggested. The q-composite scheme [3] is a variation on the EG scheme where two nodes need to share at least q keys (with $q > 1$) in order to establish a secure link between them. The q-composite scheme improves resiliency against small-scale attacks as the network becomes more vulnerable to large attacks. Du et al. [5] proposed a key predistribution scheme which also improves resiliency but at the cost of increased overheads. Although these schemes somewhat improve network resiliency, they all fail to provide *perfect* resiliency against node capture attacks. Moreover, none of them enables a node to authenticate the identity of a neighbor with which it communicates. In terms of network security this is a major drawback because *node-to-node authentication* can help detect node misbehavior, and provides resistance against node replication attacks [3].

With this in mind, Chan et al. also proposed in [3] a random pairwise key predistribution scheme with the following properties: (i) Even if some nodes are captured, the secrecy of the remaining nodes is *perfectly* preserved; (ii) Unlike earlier schemes, this pairwise scheme enables both node-to-node authentication and quorum-based revocation. The distribution scheme can be implemented through the following *offline* construction: Before deployment, each of the n sensor nodes is paired (offline) with K distinct nodes which are randomly selected from amongst all other nodes. For each such pair of sensors, 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. A secure link can then be established between two nodes if at least one of them is assigned to the other, i.e., if they have at least one pairwise key in common. Precise definitions and implementation details are given in Section II.

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; this corresponds to modelling the random pairwise distribution scheme under full visibility. The main goal of this paper is to give conditions on n and K under which $\mathbb{H}(n; K)$ is a connected graph with high probability as n grows large. As in the case of the EG scheme, such conditions might provide helpful guidelines for dimensioning purposes. In the original paper of Chan et al. [3] (as in the reference [8]), the connectivity of $\mathbb{H}(n; K)$ is analyzed by *equating* it with the Erdős-Renyi graph $\mathbb{G}(n; p)$ where $p = \frac{2K}{n}$; this constraint ensures that the link prob-

abilities in the two graphs are asymptotically matched. A formal transfer of well-known connectivity results from Erdős-Renyi graphs to $\mathbb{H}(n;K)$ suggests that the parameter K should behave like $c\log n$ for some $c > \frac{1}{2}$ in order for $\mathbb{H}(n;K)$ to be connected with a probability approaching 1 for n large. With this conclusion as a point of departure, the maximum supportable networks size is then evaluated [3], [8], and it is then argued that the random pairwise key predistribution scheme is *not* scalable.

Here we revisit this issue, and show that transferring connectivity results from Erdős-Renyi graphs to $\mathbb{H}(n;K)$ leads to *misleading* conclusions. Indeed by a *direct* analysis we show the following zero-one law: With $K \geq 2$ (resp. $K = 1$), the probability that $\mathbb{H}(n;K)$ is a connected graph approaches 1 (resp. 0) as n grows large, and the desired connectivity is therefore achievable under very small values of K.

In contrast with the EG scheme and its variations, the pairwise distribution scheme produces key rings of variable size between K and $K + (n-1)$. It is easy to see that on average only $2K$ keys will be stored in the memory module of a sensor node. Also, it can shown that the *maximum* size of a key ring is on the order $\log n$ with very high probability provided $K = O(\log n)$. Such a concentration result suggests the possibility of practically turning the pairwise scheme into a scalable one. Details will be worked out somewhere else.

The rest of the paper is organized as follows: In Section II we give a formal model for the random pairwise distribution scheme of Chan et al. The main results of the paper are given in Section III while Sections IV-V are devoted to proving the main results.

II. A PAIRWISE SCHEME

A possible offline implementation model for a random-pairwise key distribution scheme is as follows: The model is parametrized by two positive integers n and K such that $K < n$. There are n nodes, labelled $i = 1, \ldots, n$, and Kn distinct (cryptographic) keys, labelled $k = 1, \ldots, Kn$. We organize these keys into n distinct subsets of K keys, namely $\mathcal{K}_1, \ldots, \mathcal{K}_n$, with

$$\mathcal{K}_i := \{(i-1)K + 1, \ldots, iK\}, \quad i = 1, \ldots, n. \tag{1}$$

In practice the pool size P is assumed to be very large with $nK \ll P$, and can be safely taken to be infinite for the purpose of our discussion; it is not a model parameter (as was the case for the EG scheme). Also the labelling in (1) is used only for sake of concreteness, and can be replaced by any labelling which ensures that $\mathcal{K}_1, \ldots, \mathcal{K}_n$ is a partition of $\{1, \ldots, Kn\}$.

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 Γ_i of K distinct nodes selected at random from $\mathcal{N}_{-i}$. Thus, for any subset $A \subseteq \mathcal{N}_{-i}$,

$$\mathbb{P}[\Gamma_i = A] = \begin{cases} \binom{n-1}{K}^{-1} & \text{if } |A| = K \\ 0 & \text{otherwise.} \end{cases}$$

This reflects the fact that Γ_i is selected uniformly amongst all subsets of $\mathcal{N}_{-i}$ which are of size exactly K. Throughout we assume the rvs $\Gamma_1, \ldots, \Gamma_n$ to be mutually independent so that

$$\mathbb{P}[\Gamma_i = A_i,\ i = 1, \ldots, n] = \prod_{i=1}^{n} \mathbb{P}[\Gamma_i = A_i]$$

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

We now construct *offline* the key rings $\Sigma_1(n;K), \ldots, \Sigma_n(n;K)$ as follows: For each $i = 1, \ldots, n$, node i shares *each* of the K keys in $\mathcal{K}_i$ with *exactly* one of the nodes in Γ_i under the constraint that no node in Γ_i can receive more than one key from node i – A simple way to do so is to assign the K keys labelled $(i-1)K + 1, \ldots, iK$ to the K nodes in Γ_i according to increasing labels. Thus, if the random set Γ_i is realized as $\{j_1, \ldots, j_K\}$ with $1 \leq j_1 < \ldots < j_K \leq n$, then for each $k = 1, \ldots, K$, key $(i-1)K + k$ is shared with node j_k. Inherent to such an assignment is the fact that the key $(i-1)K + k$ is assigned *exclusively* to the pair of nodes i and j_k, hence the terminology pairwise distribution scheme. Other assignments are possible, e.g., according to decreasing labels or according to a random permutation.

Here we only consider the case when the sensor nodes $1, \ldots, n$ are all deployed at the same time. However, in practice the initially deployed network may have fewer than n nodes. In that case only a subset of $\{1, \ldots n\}$ will be deployed initially and the remaining sensor labels will be used at a later time if additional nodes are to be deployed. The implementation details and results regarding the case where the network is deployed *gradually* can be found in [17].

III. MAIN RESULTS AND DISCUSSION

The pairwise distribution scheme described in Section II naturally gives rise to the following class of random graphs: With $n = 2, 3, \ldots$ and positive integer K, 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

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

Let $\mathbb{H}(n;K)$ denote the undirected random graph on the vertex set $\{1, \ldots, n\}$ induced by the adjacency notion (2). Throughout let $P(n;K)$ denote the probability that $\mathbb{H}(n;K)$ is connected, namely

$$P(n;K) := \mathbb{P}[\mathbb{H}(n;K) \text{ is connected}].$$

Theorem 3.1: With any positive integer K, it holds that

$$\lim_{n\to\infty} P(n;K) = \begin{cases} 0 & \text{if } K = 1 \\ 1 & \text{if } K \geq 2. \end{cases} \tag{3}$$

We discuss the one-law and zero-law of Theorem 3.1 in Section IV and Section V, respectively. To establish the one-law we follow a variation on the approach used for proving a similar result in Erdős-Rényi graphs – We start with a union bound argument to obtain a bound on the probability that $\mathbb{H}(n;K)$ is not connected, and then show that this bound

approaches zero as n gets large. The requisite bound, also discussed in Section IV, can be given in the following simple form.

Theorem 3.2: With any positive integer $K \geq 2$, we have the bound

$$P(n;K) \geq 1 - \frac{27}{2n^2} \tag{4}$$

for all $n = 2, 3, \ldots$ sufficiently large.

The one-law in Theorem 3.1 is a clear consequence of the bound (4) which provides a rate of convergence. Analogous bounds, which are sharper than (4), are also available with an explicit dependence on K; they are omitted in the interest of brevity. These additional bounds show that the convergence in the one-law of Theorem 3.1 becomes faster with larger K, in fact as fast as $1 - O(n^{-K})$, as would be expected.

Theorem 3.1 easily yields the behavior of graph connectivity as the parameter K is scaled with n. First some terminology: We refer to any mapping $K : \mathbb{N}_0 \to \mathbb{N}_0$ as a *scaling* provided it satisfies the natural conditions

$$K_n < n, \quad n = 1, 2, \ldots . \tag{5}$$

Corollary 3.3: For any scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$, we have

$$\lim_{n\to\infty} P(n;K_n) = 1 \tag{6}$$

provided $K_n \geq 2$ for all n sufficiently large.

We stress that $\mathbb{H}(n;K)$ cannot be equated with an Erdős-Renyi graph, hence neither Theorem 3.2 nor Corollary 3.3 are consequences of classical results for Erdős-Renyi graphs [2]. To drive the point further, note the following: In many known classes of random graphs, the absence of isolated nodes and graph connectivity are asymptotically equivalent properties, e.g., Erdős-Rényi graphs [2], geometric random graphs [10] and random key graphs [12], [14]. This equivalence, when it holds, is used to advantage by first establishing the zero-one law for the absence of isolated nodes, a step which is usually much simpler to complete with the help of the method of first and second moments [9, p. 55]. However, there are no isolated in $\mathbb{H}(n;K)$ since each node is of degree at least K. Thus, the class of random graphs studied here provides an example where graph connectivity and the absence of isolated nodes are not asymptotically equivalent properties; in fact this is what makes the proof of the zero-law more intricate.

IV. A PROOF OF THE ONE-LAW

Fix $n = 2, 3, \ldots$ and consider a positive integer K. The conditions

$$2 \leq K \quad \text{and} \quad 2(K+1) < n \tag{7}$$

are assumed enforced throughout; the second condition is made to avoid degenerate situations which have no bearing on the final result. There is no loss of generality in doing so as we eventually let n go to infinity.

For any non-empty subset S of nodes, i.e., $S \subseteq \{1, \ldots, n\}$, we define the graph $\mathbb{H}(n;K)(S)$ (with vertex set S) as the subgraph of $\mathbb{H}(n;K)$ restricted to the nodes in S. We say that S is *isolated* in $\mathbb{H}(n;K)$ if there are no edges (in $\mathbb{H}(n;K)$) between the nodes in S and the nodes in the complement $S^c = \{1, \ldots, n\} - S$. This is characterized by the event $B_n(K;S)$, i.e.,

$$B_n(K;S) := [i \notin \Gamma_j, j \notin \Gamma_i,\ i \in S,\ j \in S^c] .$$

Since each node in $\mathbb{H}(n;K)$ is connected to at least K other nodes, a set S can be isolated in $\mathbb{H}(n;K)$ only if $|S| \geq K+1$.

Also, we let $C_n(K;S)$ denote the event that the induced subgraph $\mathbb{H}(n;K)(S)$ is itself connected, and write $C_n(K)$ to denote the event that $\mathbb{H}(n;K)$ is connected. Finally, we set

$$A_n(K;S) := C_n(K;S) \cap B_n(K;S).$$

The discussion starts with the following basic observation: If $\mathbb{H}(n;K)$ is *not* connected, then there must exist a subset S of nodes with $|S| \geq K+1$ such that $\mathbb{H}(n;K)(S)$ is connected while S is isolated in $\mathbb{H}(n;K)$. This is captured by the inclusion

$$C_n(K)^c \subseteq \cup_{S\in\mathcal{P}_n:\ |S|\geq K+1} A_n(K;S) \tag{8}$$

with $\mathcal{P}_n$ denoting the collection of all non-empty subsets of $\{1, \ldots, n\}$. A moment of reflection should convince the reader that this union need only be taken over all subsets S of $\{1, \ldots, n\}$ with $K+1 \leq |S| \leq \lfloor \frac{n}{2} \rfloor$. Then, a standard union bound argument immediately gives

$$\begin{aligned} \mathbb{P}[C_n(K)^c] &\leq \sum_{S\in\mathcal{P}_n: K+1\leq|S|\leq\lfloor\frac{n}{2}\rfloor} \mathbb{P}[A_n(K;S)] \\ &= \sum_{r=K+1}^{\lfloor\frac{n}{2}\rfloor} \left(\sum_{S\in\mathcal{P}_{n,r}} \mathbb{P}[A_n(K;S)] \right) \end{aligned} \tag{9}$$

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

For each $r = 1, \ldots, n$, we simplify the notation by writing $A_{n,r}(K) := A_n(K;\{1, \ldots, r\})$, $B_{n,r}(K) := B_n(K;\{1, \ldots, r\})$ and $C_{n,r}(K) := C_n(K;\{1, \ldots, r\})$. For $r = n$, the notation $C_{n,n}(K)$ coincides with $C_n(K)$ as defined earlier. Under the enforced assumptions, it is a simple matter to check by exchangeability that

$$\mathbb{P}[A_n(K;S)] = \mathbb{P}[A_{n,r}(K)], \quad S \in \mathcal{P}_{n,r}$$

and the expression

$$\sum_{S\in\mathcal{P}_{n,r}} \mathbb{P}[A_n(K;S)] = \binom{n}{r} \mathbb{P}[A_{n,r}(K)] \tag{10}$$

follows since $|\mathcal{P}_{n,r}| = \binom{n}{r}$. Substituting into (9) we obtain the bounds

$$\mathbb{P}[C_n(K)^c] \leq \sum_{r=K+1}^{\lfloor\frac{n}{2}\rfloor} \binom{n}{r} \mathbb{P}[B_{n,r}(K)] \tag{11}$$

as we note the inclusion $A_{n,r}(K) \subseteq B_{n,r}(K)$.

For each $r = K+1, \ldots, n$, it is easy to check that

$$\mathbb{P}[B_{n,r}(K)] = \left(\frac{\binom{r-1}{K}}{\binom{n-1}{K}} \right)^r \cdot \left(\frac{\binom{n-r-1}{K}}{\binom{n-1}{K}} \right)^{n-r} . \tag{12}$$

Therefore, reporting (12) into (11) we get

$$\mathbb{P}\left[C_n(K)^c\right] \leq \sum_{r=K+1}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{r} \left(\frac{\binom{r-1}{K}}{\binom{n-1}{K}} \right)^r \left(\frac{\binom{n-r-1}{K}}{\binom{n-1}{K}} \right)^{n-r}. \quad (13)$$

For $0 \leq K \leq x \leq y$, we have

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

since $\frac{x-\ell}{y-\ell}$ decreases as ℓ increases from $\ell = 0$ to $\ell = K-1$. Reporting this fact into (13) and using the standard bound

$$\binom{n}{r} \leq \left(\frac{ne}{r} \right)^r, \quad r = 1, \ldots, n$$

we conclude that

$$\begin{aligned}
&\mathbb{P}\left[C_n(K)^c\right] \\
&\leq \sum_{r=K+1}^{\lfloor \frac{n}{2} \rfloor} \left(\frac{ne}{r} \right)^r \left(\frac{r-1}{n-1} \right)^{rK} \left(1 - \frac{r}{n-1} \right)^{K(n-r)} \\
&\leq \sum_{r=K+1}^{\lfloor \frac{n}{2} \rfloor} \left(\frac{ne}{r} \right)^r \left(\frac{r}{n} \right)^{rK} \left(1 - \frac{r}{n} \right)^{K(n-r)} \\
&\leq \sum_{r=K+1}^{\lfloor \frac{n}{2} \rfloor} \left(\frac{ne}{r} \right)^r \left(\frac{r}{n} \right)^{rK} e^{-rK\frac{(n-r)}{n}} \\
&= \sum_{r=K+1}^{\lfloor \frac{n}{2} \rfloor} \left(\left(\frac{r}{n} \right)^{K-1} e^{1-K\frac{(n-r)}{n}} \right)^r. \quad (14)
\end{aligned}$$

On the range $r = K+1, \ldots, \lfloor \frac{n}{2} \rfloor$ with $K \geq 2$, we have

$$K\ \frac{n-r}{n} \geq K\ \frac{n - \lfloor \frac{n}{2} \rfloor}{n} \geq \frac{K}{2} \geq 1,$$

whence

$$e^{1-K\frac{(n-r)}{n}} \leq 1.$$

On the same range we also have

$$\left(\frac{r}{n} \right)^{K-1} \leq \frac{r}{n}$$

since $K \geq 2$. Reporting these facts into (14) we find

$$\mathbb{P}\left[C_n(K)^c\right] \leq \sum_{r=K+1}^{\lfloor \frac{n}{2} \rfloor} \left(\frac{r}{n} \right)^r \leq \sum_{r=3}^{\lfloor \frac{n}{2} \rfloor} \left(\frac{r}{n} \right)^r. \quad (15)$$

With the help of (15) it is already possible to show that $\lim_{n\to\infty} \mathbb{P}\left[C_n(K)^c\right] = 0$, hence (3), by an easy application of the Bounded Convergence Theorem. Details are omitted in the interest of brevity.

However, the bounds in (15) do pave the way for the more compact bound of Theorem 3.2. This is a consequence of the equality

$$\max\left(\left(\frac{r}{n} \right)^r :\ r = 3, \ldots, \left\lfloor \frac{n}{2} \right\rfloor \right) = \left(\frac{3}{n} \right)^3 \quad (16)$$

valid for all n sufficiently large. This fact is easily validated by standard calculus arguments. Now, putting (16) into (15) yields

$$\mathbb{P}\left[C_n(K)^c\right] \leq \sum_{r=3}^{\lfloor \frac{n}{2} \rfloor} \left(\frac{3}{n} \right)^3$$

for all n sufficiently large, and the bound (4) is established.

V. A PROOF OF THE ZERO-LAW

First some terminology: When $K = 1$, the random sets $\Gamma_1, \ldots, \Gamma_n$ are now singletons, and can be interpreted as $\{1, \ldots, n\}$-valued rvs (as we do from now on) such that $\Gamma_i \neq i$ for each $i = 1, \ldots, n$. Thus, Γ_i is the node selected at random which becomes associated (paired) with node i.

With this in mind, a *formation* is any sequence $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_n)$ such that for each $i = 1, \ldots, n$, the component γ_i is an element of $\{1, \ldots, n\}$ such that $\gamma_i \neq i$. In other words, $\boldsymbol{\gamma}$ is one of the $(n-1)^n$ possible realizations of the rvs $(\Gamma_1, \ldots, \Gamma_n)$.

With any formation $\boldsymbol{\gamma}$ we associate a *directed* graph on the vertex set $\{1, \ldots, n\}$ in an obvious manner: There is a directed edge from node i to node j if $\gamma_i = j$. This directed graph is denoted by $H_{\boldsymbol{\gamma}}(n)$. As there are $(n-1)^n$ possible formations, there are $(n-1)^n$ distinct directed graphs so defined. Under the pairwise distribution scheme considered here, each of these graphs is equally likely, so that we have

$$P(n;1) = \frac{\sum_{\boldsymbol{\gamma}} \mathbf{1}\left[H_{\boldsymbol{\gamma}}(n) \text{ is connected } \right]}{(n-1)^n} \quad (17)$$

where the summation $\sum_{\boldsymbol{\gamma}}$ is taken over all possible formations. Here, we have used the conventional notion of connectivity for directed graphs: A directed graph is connected if and only if the underlying *undirected* graph is connected – This is to be distinguished from the notion of strong connectivity defined for directed graphs. The desired zero-law will be established if we can show that

$$\lim_{n\to\infty} \frac{\sum_{\boldsymbol{\gamma}} \mathbf{1}\left[H_{\boldsymbol{\gamma}}(n) \text{ is connected } \right]}{(n-1)^n} = 0. \quad (18)$$

From now on, let $H^\star_{\boldsymbol{\gamma}}(n)$ denote the underlying undirected graph of $H_{\boldsymbol{\gamma}}(n)$. We note that $H^\star_{\boldsymbol{\gamma}}(n)$ is a realization of the random graph $\mathbb{H}(n;1)$ when $(\Gamma_1, \ldots, \Gamma_n) = \boldsymbol{\gamma}$. For each formation $\boldsymbol{\gamma}$, we can easily validate the following observations:

1) By definition, $H_{\boldsymbol{\gamma}}(n)$ is connected if and only if $H^\star_{\boldsymbol{\gamma}}(n)$ is connected.
2) Since $H_{\boldsymbol{\gamma}}(n)$ has n directed edges, the undirected graph $H^\star_{\boldsymbol{\gamma}}(n)$ can have *at most* n edges.
3) If $H_{\boldsymbol{\gamma}}(n)$ is connected, then $H^\star_{\boldsymbol{\gamma}}(n)$ should have *at least* $n-1$ edges.
4) If $H_{\boldsymbol{\gamma}}(n)$ is connected and $H^\star_{\boldsymbol{\gamma}}(n)$ has $n-1$ edges, then $H_{\boldsymbol{\gamma}}(n)$ has exactly one bi-directional edge.
5) If $H_{\boldsymbol{\gamma}}(n)$ is connected and has (exactly) one bi-directional edge, then $H^\star_{\boldsymbol{\gamma}}(n)$ is a necessarily a *tree* with $n-1$ edges.
6) If $H_{\boldsymbol{\gamma}}(n)$ is connected and $H^\star_{\boldsymbol{\gamma}}(n)$ has n edges, then $H_{\boldsymbol{\gamma}}(n)$ has exactly one *cycle*.

Case I – $\mathbb{H}(n;1)$ is connected and has $n-1$ edges: Thus, $\mathbb{H}(n;1)$ is a tree. With $\mathcal{T}_n$ denoting the collection of labelled trees on the set of vertices $\{1,\ldots,n\}$, we have $|\mathcal{T}_n| = n^{n-2}$. by Cayley's formula. Noting also that a given tree is the underlying undirected graph for $n-1$ different formations (corresponding to $n-1$ possible places for the single bi-directional edge), we get

$$\begin{aligned}
&\mathbb{P}\left[\mathbb{H}(n;1) \text{ is connected and has } n-1 \text{ edges}\right] \\
&= \frac{1}{(n-1)^n} \cdot \sum\nolimits_{\boldsymbol{\gamma}} \mathbf{1}\left[\begin{array}{c} H_{\boldsymbol{\gamma}}(n) \text{ is connected and} \\ \text{has one bi-directional edge} \end{array}\right] \\
&= \frac{1}{(n-1)^n} \cdot \sum\nolimits_{\boldsymbol{\gamma}} \sum_{T \in \mathcal{T}_n} \mathbf{1}\left[H^{\star}_{\boldsymbol{\gamma}}(n) = T\right] \\
&= \frac{1}{(n-1)^n} \cdot (n-1) \cdot n^{n-2} \\
&= \frac{1}{n} \cdot \left(\frac{n}{n-1}\right)^{n-1}. \qquad (19)
\end{aligned}$$

It is now clear that

$$\lim_{n\to\infty} \mathbb{P}\left[\begin{array}{c} \mathbb{H}(n;1) \text{ is connected} \\ \text{and has } n-1 \text{ edges} \end{array}\right] = 0. \qquad (20)$$

Case II – $\mathbb{H}(n;1)$ is connected and has n edges: This corresponds to all formations $\boldsymbol{\gamma}$ such that $H^{\star}_{\boldsymbol{\gamma}}(n)$ is connected and has exactly one cycle. It is not difficult to see that a connected graph with only one cycle can be the underlying undirected graph for two different formations (corresponding to the two possible orientations of the cycle). For instance, consider a connected graph on n nodes with exactly one cycle. This graph necessarily has n edges and therefore the original directed graph $H_{\boldsymbol{\gamma}}(n)$ cannot have a bi-directional edge. Without loss of generality, assume that the cycle consists of nodes 1, 2, 3, 4 with edges $1 \sim 2, 2 \sim 3, 3 \sim 4, 4 \sim 1$. Then the two possible formations are $\{2,3,4,1,\gamma_5,\gamma_6,\ldots\gamma_n\}$ and $\{4,1,2,3,\gamma_5,\gamma_6,\ldots\gamma_n\}$. Similar arguments can be made for all possible cycles. Since there can be no other cycles or bi-directional edges in the rest of the graph, these two formations will be the only ones that give rise to that particular undirected structure.

Now let $\mathcal{T}_n^+$ denote the set of undirected graphs on n nodes which are connected and have exactly n edges. We find

$$\begin{aligned}
&\mathbb{P}\left[\mathbb{H}(n;1) \text{ is connected and has } n \text{ edges}\right] \\
&= \frac{1}{(n-1)^n} \cdot \sum\nolimits_{\boldsymbol{\gamma}} \mathbf{1}\left[\begin{array}{c} H_{\boldsymbol{\gamma}}(n) \text{ is connected and} \\ \text{has exactly one cycle} \end{array}\right] \\
&= \frac{1}{(n-1)^n} \cdot \sum\nolimits_{\boldsymbol{\gamma}} \sum_{G \in \mathcal{T}_n^+} \mathbf{1}\left[H^{\star}_{\boldsymbol{\gamma}}(n) = G\right] \\
&= \frac{1}{(n-1)^n} \cdot 2 \cdot |\mathcal{T}_n^+|. \qquad (21)
\end{aligned}$$

However, it is known [7, p. 133-134] that

$$|\mathcal{T}_n^+| \sim \frac{1}{4}\sqrt{2\pi} n^{n-\frac{1}{2}},$$

and reporting this fact into (21) gives

$$\begin{aligned}
&\mathbb{P}\left[\mathbb{H}(n;1) \text{ is connected and has } n \text{ edges}\right] \\
&\sim \frac{\sqrt{2\pi}}{2}\left(\frac{n}{n-1}\right)^n n^{-\frac{1}{2}} \\
&\sim \frac{\sqrt{2\pi}e}{2} n^{-\frac{1}{2}}. \qquad (22)
\end{aligned}$$

It is now immediate that

$$\lim_{n\to\infty} \mathbb{P}\left[\mathbb{H}(n;1) \text{ is connected and has } n \text{ edges}\right] = 0.$$

Together with (20) and Facts 2-3, we now conclude that (18) holds.

REFERENCES

[1] S.R. Blackburn and S. Gerke, "Connectivity of the uniform random intersection graph," *Discrete Mathematics* **309** (2009), pp. 5130-5140.

[2] B. Bollobás, *Random Graphs*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge (UK), 2001.

[3] H. Chan, A. Perrig, D. Song, "Random Key Predistribution Schemes for Sensor Networks," Proceedings of the 2003 IEEE Symposium on Research in Security and Privacy (SP 2003), Oakland (CA), May 2003, pp. 197-213.

[4] R. Di Pietro, L.V. Mancini, A. Mei, A. Panconesi and J. Radhakrishnan, "Redoubtable sensor networks," *ACM Transactions on Information Systems Security* **TISSEC 11** (2008), pp. 1-22.

[5] W. Du, J. Deng, Y.S. Han and P.K. Varshney, "A pairwise key predistribution scheme for wireless sensor networks," in Proceedings of the 10th ACM Conference on Computer and Communications Security (CCS 2003), Washington (DC), October 2003, pp. 42-51.

[6] L. Eschenauer and V.D. Gligor, "A key-management scheme for distributed sensor networks," in Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS 2002), Washington (DC), November 2002, pp. 41-47.

[7] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge (UK), January 2009.

[8] J. Hwang and Y. Kim, "Revisiting random key pre-distribution schemes for wireless sensor networks," in the Proceedings of the Second ACM Workshop on Security of Ad Hoc And Sensor Networks (SASN 2004), Washington (DC), October 2004.

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

[10] M.D. Penrose, *Random Geometric Graphs*, Oxford Studies in Probability **5**, Oxford University Press, New York (NY), 2003.

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

[12] K. Rybarczyk "Diameter, connectivity and phase transition of the uniform random intersection graph," Submitted to *Discrete Mathematics*, July 2009.

[13] D.-M. Sun and B. He, "Review of key management mechanisms in wireless sensor networks," *Acta Automatica Sinica* **12** (2006), pp. 900-906.

[14] O. Yağan and A.M. Makowski, "On the random graph induced by a random key predistribution scheme under full visibility," In Proceedings of the IEEE International Symposium on Information Theory (ISIT 2008), Toronto (ON), June 2008.

[15] O. Yağan and A.M. Makowski, "Connectivity results for random key graphs," In Proceedings of the IEEE International Symposium on Information Theory, (ISIT 2009), Seoul (S. Korea), June 2009.

[16] O. Yağan and A.M. Makowski, "Zero-one laws for connectivity in random key graphs," Submitted to *IEEE Trnasactions on Information Theory*, January 2009. Available online at arXiv:0908.3644v1 [math.CO], August 2009. Earlier draft available online at http://www.lib.umd.edu/drum/handle/1903/8716, January 2009.

[17] O. Yağan and A. M. Makowski, "Modelling random pairwise key distribution schemes – Gradual deployment," submitted for inclusion in the program of WiOpt 2010, Avignon (France), June 2010.