On the Size of the Giant Component in Inhomogeneous Random K-out Graphs

Mansi Sood and Osman Yağan

***Abstract*— Inhomogeneous random K-out graphs were recently introduced to model heterogeneous sensor networks secured by random pairwise key predistribution schemes. First, each of the n nodes is classified as type-1 (respectively, type-2) with probability $0 < \mu < 1$ (respectively, $1 - \mu$) independently from each other. Next, each type-1 (respectively, type-2) node draws 1 *arc* towards a node (respectively, K_n arcs towards K_n distinct nodes) selected uniformly at random, and then the orientation of the arcs is ignored. It was recently established that this graph, denoted by $\mathbb{H}(n; \mu, K_n)$, is connected *with high probability* (whp) if and only if $K_n = \omega(1)$. In other words, if $K_n = O(1)$, then $\mathbb{H}(n; \mu, K_n)$ has a positive probability of being *not* connected as n gets large. Here, we study the *size* of the largest connected *subgraph* of $\mathbb{H}(n; \mu, K_n)$ when $K_n = O(1)$. We show that the trivial condition of $K_n \geq 2$ for all n is sufficient to ensure that inhomogeneous K-out graph has a connected component of size $n - O(1)$ whp. Put differently, even with $K_n = 2$, all but finitely many nodes will form a connected sub-network in this model under any $0 < \mu < 1$. We present an upper bound on the probability that more than M nodes are outside of the largest component, and show that this decays as $O(1) \exp\{-M(1 - \mu)(K_n - 1)\} + o(1)$. Numerical results are presented to demonstrate the size of the largest connected component when the number of nodes is finite.**

I. Introduction

Random graph modeling is as an important framework for developing fundamental insights into the structure and dynamics of several complex real-world networks including social networks, economic networks and communication networks [1]–[4]. In the context of wireless sensor networks (WSNs), random graph models have been used widely [5], [6] in the design and performance evaluation of *random key predistribution schemes* that were proposed for ensuring their secure connectivity [5], [7], [8].

The *random K-out graph* is one of the earliest models studied in the literature [9], [10]. Denoted here by $\mathbb{H}(n; K)$, it is constructed as follows. Each of the n nodes draws K arcs towards K distinct nodes chosen uniformly at random among all others. The orientation of the arcs is then ignored, yielding an *undirected* graph. Random K-out graphs have received renewed interest for analyzing secure sensor networks and anonymous routing in cryptocurrency networks. In the context of wireless sensor networks, random K-out graphs have been studied [11]–[14] to model the random *pairwise* key predistribution scheme [15]. Along with the original key predistribution scheme proposed by Escheanuer and Gligor [5], the pairwise scheme is one of the most widely recognized security protocols for WSNs. Cryptocurrency networks provide another application where a structure similar to random K-out graphs has been proposed to make message propagation robust to *de-anonymization* attacks [16, Algorithm 1].

In recent years, the analysis of heterogeneous variants of classical random graph models has emerged as an important topic [17]–[22], owing to the fact that real-world network applications are increasingly *heterogeneous* with participating nodes having different capabilities and (security and connectivity) requirements [1], [23]–[26]. A heterogeneous variant of random K-out graph, known as the *inhomogeneous random K-out graph*, was proposed recently to model networks secured by *heterogeneous* random *pairwise* key predistribution schemes [17], [18]. In the inhomogeneous random K-out graph, each node is independently classified as type-1 (respectively, type-2) with probability μ (respectively, $1 - \mu$). Then, each type-1 (respectively, type-2) node selects one node (respectively, $K_n \geq 2$ nodes) uniformly at random from all other nodes; see Figure 1. The notation K_n indicates that the number of selections made by type-2 nodes *scales* as a function of the number of nodes n. We denote the inhomogeneous random K-out graph on n nodes with parameters μ and K_n as $\mathbb{H}(n; \mu, K_n)$.

In several real-world networks, *connectivity* is attributed to be a fundamental determinant of system performance. It was established in [9], [11] that random K-out graphs are connected (respectively, not connected) *with high probability* (whp) when $K \geq 2$ (respectively, when $K = 1$); i.e.,

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

In [17], it was shown that for any $0 < \mu < 1$, the inhomogeneous random K-out graph is connected whp if and only if K_n grows unboundedly large with n; i.e.,

$$\lim_{n\to\infty} \mathbb{P}\left[\mathbb{H}(n; \mu, K_n) \text{ is connected}\right] = \begin{cases} 1 & \text{if } K_n \to \infty \\ < 1 & \text{otherwise.} \end{cases} \tag{2}$$

As seen from (2), ensuring connectivity of $\mathbb{H}(n; \mu, K_n)$ requires $K_n = \omega(1)$. Although it is desirable to have a connected network, in several practical applications, resource constraints can potentially limit the number of links that can be successfully established. For instance, if the power available for transmission is limited, the underlying physical network may not be dense enough to guarantee global connectivity with key predistribution schemes [27]. Depending on the nature of the application, it may suffice to have a large

M. Sood and O. Yağan are with Department of Electrical and Computer Engineering and CyLab, Carnegie Mellon University, Pittsburgh, PA, 15213 USA. Email: {msood@andrew.cmu.edu, oyagan@ece.cmu.edu}

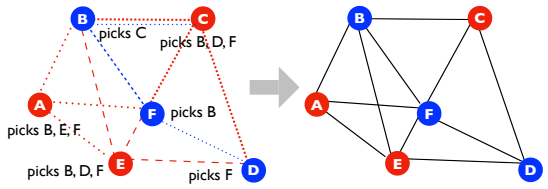

Fig. 1. An inhomogeneous random K-out graph with 6 nodes. Nodes A, C and E are type-2 and the rest (B, D, F) are type-1. Each type-1 (resp. type-2) node selects 1 (resp. $K_n = 3$) node uniformly at random. An edge is drawn between two nodes if at least one selects the other.

connected sub-network spanning almost the entire network [28]. For example, if a sensor network is designed to monitor temperature of a field, it may suffice to aggregate readings from a majority of sensors in the field [29].

With this in mind, the question which we address here is when K_n is bounded (i.e., $K_n = O(1)$), how many nodes are contained in the largest connected sub-network (i.e., component) of $\mathbb{H}(n; \mu, K_n)$? In the literature on random graphs, this is often studied in terms of the *existence* and size of the *giant component*, defined as a connected sub-network comprising $\Omega(n)$ nodes; see [30] for a classical study on the size of the giant component of Erdős-Rényi graphs.

We establish that the inhomogeneous random K-out graph contains a giant component as long as the trivial conditions $0 < \mu < 1$ and $K_n \geq 2$ (for all n) hold. In fact, we show under the same conditions that the graph contains a connected sub-network of size $n - O(1)$ whp. Put differently, *all but finitely many* nodes will be contained in the giant component of $\mathbb{H}(n; \mu, K_n)$, as n goes to infinity. This is also demonstrated through numerical experiments where we observe that with $n = 5000, \mu = 0.9, K_n = 2$, at most 45 nodes turned out to be outside the largest connected component across 100,000 experiments; see Section III for details. Our main result follows from an upper bound on the probability that more than M nodes are outside of the giant component. We show that this probability decays at least as fast as $O(1) \exp\{-M(1-\mu)(K_n - 1)\} + o(1)$ providing a clear trade-off between K_n and the fraction $(1-\mu)$ of nodes that make K_n selections.

We close by describing a potential application of (inhomogeneous) random K-out graphs. Given their sparse yet connected structure, these graphs can be useful for analyzing payment channel networks (PCNs) wherein edges represent the funds escrowed in a bidirectional overlay network on top of the cryptocurrency network [31]. Recent work in the realm of cryptocurrency networks has closely looked at the topological properties of PCNs and their impact on the achieved throughput [32]–[34]. A key aspect of PCNs is the trade-off between the number of edges in the network (which is constrained since funds need to be committed on each edge) and its connectivity (which is desirable so that any pair of nodes can perform transactions with each other). The results established here show that the construction of inhomogeneous random K-out graphs leads to almost all nodes being connected with each other (as part of the largest connected component) with relatively small number of edges per node. For instance, with $K = 2$ and $\mu = 0.5$, each node will have 3 edges on average. In fact, the Lightning Network dataset from December 2018 shows that it contains 2273 nodes, of which 2266 are contained in the largest connected component while the remaining 7 nodes being in three isolated components.

All limits are understood with the number of nodes n going to infinity. While comparing asymptotic behavior of a pair of sequences $\{a_n\}, \{b_n\}$, we use $a_n = o(b_n)$, $a_n = \omega(b_n)$, $a_n = O(b_n)$, $a_n = \Theta(b_n)$, and $a_n = \Omega(b_n)$ with their meaning in the standard Landau notation. All random variables are 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 $\mathbb{E}$. For an event A, its complement is denoted by A^c. We say that an event occurs with high probability (whp) if it holds with *probability tending to one* as $n \to \infty$. We denote the cardinality of a discrete set A by $|A|$ and the set of all positive integers by $\mathbb{N}_0$. For events A and B, we use $A \implies B$ with the meaning that $A \subseteq B$.

II. Inhomogeneous Random K-out Graph

Let $\mathcal{N} := \{1, 2, \ldots, n\}$ denote the set of vertex labels and let $\mathcal{N}_{-i} := \{1, 2, \ldots, n\} \setminus i$. In its simplest form, the inhomogeneous random K-out graph is constructed on the vertex set $\{v_1, \ldots, v_n\}$ as follows. First, each vertex is assigned as type-1 (respectively, type-2) with probability μ (respectively, $1 - \mu$) independently from other nodes, where $0 < \mu < 1$. Next, each type-1 (respectively, type-2) node selects K_1 (respectively, K_2) distinct nodes uniformly at random among all other nodes. For each $i \in \mathcal{N}$, let $\Gamma_{n,i} \subseteq \mathcal{N}_{-i}$ denote the labels corresponding to the selections made by v_i. Under the aforementioned assumptions, $\Gamma_{n,1}, \ldots, \Gamma_{n,n}$ are mutually independent *given* the types of nodes. We say that distinct nodes v_i and v_j are adjacent, denoted by $v_i \sim v_j$ if at least one of them picks the other. Namely,

$$v_i \sim v_j \qquad \text{if} \qquad j \in \Gamma_{n,i} \ \vee \ i \in \Gamma_{n,j}. \tag{3}$$

The inhomogeneous random K-out graph is then defined on the vertices $\{v_1, \ldots, v_n\}$ through the adjacency condition (3). More general constructions with an arbitrary number of node types is also possible [18], and the implications of our results for such cases is presented in the Appendix.

Without loss of generality, we assume that $1 \leq K_1 < K_2$. From (1), it can be seen that the inhomogeneous random K-out graph will be connected whp if $K_1 \geq 2$. Therefore, interesting cases arise for the connectivity and size of the largest component only when $K_1 = 1$; i.e., when each node has a positive probability μ of selecting only one other node. As in [17], we thus assume that $K_1 = 1$ which in turn implies that $K_2 \geq 2$. For generality, we let K_2 to scale with (i.e., to be a function of) n and simplify the notation by denoting the corresponding mapping as K_n. Put differently, we consider the inhomogeneous random K-out graph, denoted as $\mathbb{H}(n; \mu, K_n)$, where each of the n nodes selects one other node with probability $0 < \mu < 1$ and K_n other nodes with probability $1 - \mu$; the edges are

then constructed according to (3). Throughout, we assume that $K_n \geq 2$ for all n in line with the assumption that $K_2 > K_1 = 1$. We denote the average number of selections made by each node in $\mathbb{H}(n;\mu,K_n)$ by $\langle K_n \rangle$. Observe that

$$\langle K_n \rangle = \mu + (1-\mu)K_n. \tag{4}$$

III. Main Results and Discussion

A. The Main Result

It is known from [17] that $\mathbb{H}(n;\mu,K_n)$ is connected whp *only if* $K_n = \omega(1)$. A natural question is then to ask what would happen if K_n is *bounded*, i.e, when $K_n = O(1)$. It was shown, again in [17], that $\mathbb{H}(n;\mu,K_n)$ has a positive probability of being *not* connected in that case. Thus, it is of interest to analyze whether the network has a connected sub-network containing a *large* number of nodes, or it consists merely of *small* sub-networks isolated from each other. To answer this question, we formally define connected components and then state our main result characterizing the size of the largest connected component of $\mathbb{H}(n;\mu,K_n)$ when $K_n = O(1)$.

Definition 3.1 (Connected Components): Nodes v_1 and $v_2 \in \mathcal{N}$ are said to be *connected* if there exists a path of edges connecting them. The connectivity of a pair of nodes forms an equivalence relation on the set of nodes. Consequently, there is a partition of the set of nodes $\mathcal{N}$ into non-empty sets $C_1, C_2, \ldots, C_m$ (referred to as connected components) such that two vertices v_1 and v_2 are connected if and only if there exists $i \in \{1,\ldots,m\}$ for which $v_1, v_2 \in C_i$; see [35, p. 13].

In light of the above definition, a graph is connected if it consists of only one connected component. In all other cases, the graph is *not* connected and has at least two connected components that have no edges in between. It is of interest to analyze the fraction of the nodes contained in the *largest* connected component as the number of nodes grows. In particular, a graph with n nodes is said to have a *giant* component if its largest connected component is of size $\Omega(n)$.

Let $C_{\max}(n;\mu,K_n)$ denote the set of nodes in the largest connected component of $\mathbb{H}(n;\mu,K_n)$. Our main results, presented below, show that $|C_{\max}(n;\mu,K_n)| = n - O(1)$ whp. Namely, $\mathbb{H}(n;\mu,K_n)$ has a giant component that contains *all but finitely many* of the nodes whp. First, we show that the probability of at least M nodes being *outside* of $C_{\max}(n;\mu,K_n)$ decays exponentially fast with M.

Theorem 3.2: For the inhomogeneous random graph $\mathbb{H}(n;\mu,K_n)$ with $K_n \geq 2$ $\forall n$ and $K_n = O(1)$ we have for each $M = 1,2,\ldots$ that

$$\begin{aligned}&\mathbb{P}\left[|C_{\max}(n;\mu,K_n)| \leq n - M\right] \\ &\leq \frac{\exp\{-M\left(\langle K_n\rangle - 1\right)\left(1-o(1)\right)\}}{1-\exp\{-\left(\langle K_n\rangle - 1\right)\left(1-o(1)\right)\}} + o(1).\end{aligned} \tag{5}$$

The proof of Theorem 3.2 relies on the connection between the *non-existence* of sub-graphs with size exceeding M that are *isolated* from the rest of the graph, and the size of the of largest component being at least $n - M$. This approach is inspired by [28] and differs from the branching process technique typically employed in the random graph literature, e.g., in the case of Erdős-Rényi graphs [36, Ch. 4]. The proof of Theorem 3.2 is presented in Section IV.

Corollary 3.3: For the inhomogeneous random graph $\mathbb{H}(n;\mu,K_n)$ with $K_n \geq 2$ $\forall n$ and $K_n = O(1)$ we have

$$|C_{\max}(n;\mu,K_n)| = n - O(1) \text{ whp.} \tag{6}$$

Proof. Consider an arbitrary sequence $x_n = \omega(1)$. Substituting M with x_n in (5), we readily see that

$$\lim_{n\to\infty} \mathbb{P}\left[n - |C_{\max}(n;\mu,K_n)| \leq x_n\right] = 1. \tag{7}$$

Namely, we have

$$n - |C_{\max}(n;\mu,K_n)| \leq x_n \text{ whp for any } \; x_n = \omega(1). \tag{8}$$

This is equivalent to the number of nodes $(n - |C_{\max}(n;\mu,K_n)|)$ outside the largest connected component being *bounded*, i.e., $O(1)$, with high probability. This fact is sometimes stated using the probabilistic big-O notation, O_p. A random sequence $f_n = O_p(1)$ if for any $\varepsilon > 0$ there exists finite integers $M(\varepsilon)$ and $n(\varepsilon)$ such that $\mathbb{P}[f_n > M(\varepsilon)] < \varepsilon$ for all $n \geq n(\varepsilon)$. In fact, we see from [37, Lemma 3] that (8) is equivalent to having $n - |C_{\max}(n;\mu,K_n)| = O_p(1)$ Here, we equivalently state this as

$$n - |C_{\max}(n;\mu,K_n)| = O(1) \;\; \text{whp},$$

giving readily (6). ∎

We extend Corollary 3.3 to inhomogeneous random K-out graphs with *arbitrary* number of node types; see Appendix.

B. Discussion

Theorem 3.2 shows that for arbitrary $0 < \mu < 1$ and even with $K_n = 2$, the largest connected component in $\mathbb{H}(n;\mu,K_n)$ spans $n - O(1)$ nodes whp. We expect that especially in resource-constrained environments (e.g., IoT type settings), it will be advantageous to have a *large* connected component reinforcing the usefulness of the heterogeneous pairwise key predistribution scheme for ensuring secure communications in such applications; see [15], [17] for other advantages of the (heterogeneous) pairwise scheme.

It is worth emphasizing that the largest connected component of $\mathbb{H}(n;\mu,K_n)$, whose size is given in (6), is *much larger* than what is strictly required to qualify it as a *giant* component, i.e, the condition that $|C_{\max}(n;\mu,K_n)| = \Omega(n)$. In fact, for most random graph models, including Erdős-Rényi graphs [30], random key graphs [38, Theorem 2], studies on the size of the largest connected component are focused on characterizing the behavior of $|C_{\max}|/n$ as n gets large; this amounts to studying the *fractional* size of the largest connected component. Our result given at (6) goes beyond looking at the fractional size of the largest component, for which it gives $\frac{|C_{\max}(n;\mu,K_n)|}{n} \to_p 1$. This is equivalent to having $|C_{\max}(n;\mu,K_n)| = n - o(n)$. However, even having $|C_{\max}(n;\mu,K_n)| = n - o(n)$ leaves the possibility that as

many as $n^{0.99}$ nodes are *not* part of the largest connected component. Thus, our result, showing that at most $O(1)$ nodes are outside the largest connected component *whp*, is sharper than existing results on the *fractional* size of the largest connected component.

Our result highlights a major difference of inhomogeneous random K-out graphs with classical models such as Erdős-Rényi (ER) graphs [10], [30]. We provide an example to compare the size of the giant component in $\mathbb{H}(n;\mu,K_n)$ and ER graphs with the same mean degree. For $\mathbb{H}(n;\mu,K_n)$, we set $K_n=2$ and $\mu=0.9$, which yields a mean node degree of $(1-o(1))2\langle K_n\rangle \approx 2(0.9+0.1\times 2)=2.2$; see Appendix. Let $\mathbb{G}(n;p_n)$ denote the ER graph on n nodes with edge probability $p_n\in[0,1]$. We set $p_n=2.2/n$ to get a mean degree of 2.2. Thus, the mean number of edges in both these models match. It is known that for $p_n=c/n$ and $c>1$, the ER graph has a giant component of size $\beta n(1+o(1))$ whp, where $\beta\in(0,1]$ is the solution of $\beta+e^{-\beta c}=1$. Substituting $p=2.2/n$, the largest connected component of the ER graph $\mathbb{G}(n;2.2/n)$ is of size $\approx 0.8437n+o(n)$ whp. For an ER graph over 5000 nodes, this corresponds to over 700 nodes being isolated from the largest component. In contrast, Theorem 3.2 shows that the largest connected component of $\mathbb{H}(n;\mu,K_n)$ would be much larger. Namely, $C_{\max}(n;\mu,K_n)=n-O(1)$ whp. This is verified in our experiments in Figure 2, where for a network of 5000 nodes, at most 45 nodes are outside the largest connected component in 100,000 experiments.

C. Numerical Results

Through simulations, we examine the size of $C_{\max}(n;\mu,K_n)$ when the number of nodes is finite. We first explore the impact of varying the probability μ of a node being type-1 on the size of the largest connected component. We generate 100,000 independent realizations of $\mathbb{H}(n;\mu,K_n)$ with $K_n=2$ for $n=1000$ and $n=5000$, varying μ between 0.1 and 0.9 in increments of 0.1. Since, Theorem 3.2 states that the size of the largest connected component is $n-O(1)$ whp, we focus on the *minimum* size of the largest component observed in 100,000 experiments. The *average* size of the largest component is also shown for comparison in Figure 2. We see that even when the probability of a node being type-1 is as high as 0.9, setting $K_n=2$ suffices to have a connected component spanning almost all of the nodes. For $n=1000$ and 5000, at most 60 and 45 nodes, respectively, are found to be outside the largest connected component. The observation that the number of nodes outside the largest connected component does not scale with n is consistent with Theorem 3.2 and Corollary 3.3.

The next set of experiments probes the impact of varying the number K_n of edges pushed by type-2 nodes while μ is fixed. We generate 100,000 independent realizations of $\mathbb{H}(n;\mu,K_n)$ for $n=5000$ while keeping μ fixed at 0.9 and varying K_n between 2 and 10 in increments of 1. Increasing K_n has an impact similar to decreasing μ and we see in Figure 3 that both the average and the minimum size of the largest connected component increases nearly monotonically. Given that increasing K_n (or, decreasing μ) increases $\langle K_n\rangle$ in view of (4), this observation is consistent with our main result given in Theorem 3.2; i.e., with the fact that $\mathbb{P}\left[n-|C_{\max}(n;\mu,K_n)|>M\right]$ decays to zero exponentially with $(\langle K_n\rangle-1)M$.

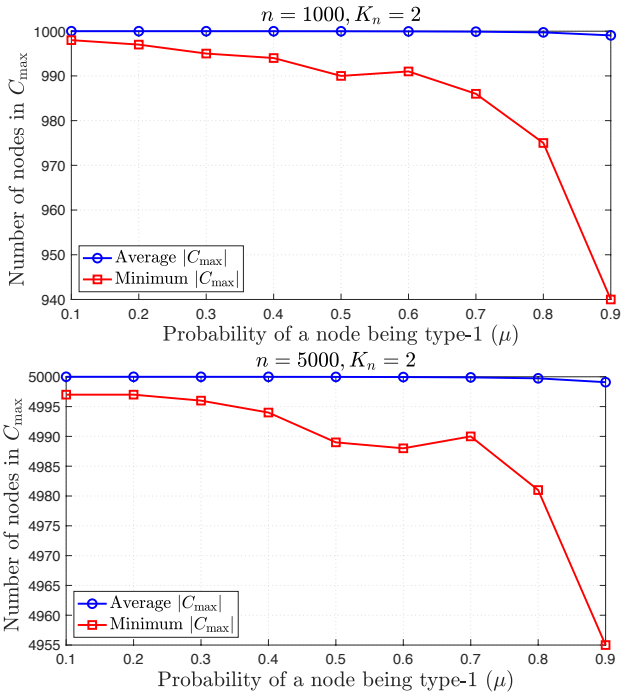

Fig. 2. *Average and minimum number of nodes contained in the largest connected component of $\mathbb{H}(n;\mu,K_n)$ with $K_n=2$, $n=1000, 5000$ and $\mu\in\{0.1,\ldots,0.9\}$. Even when $\mu=0.9$, setting $K_n=2$ is enough to ensure that almost all of the nodes form a connected component; at most 45 out of 5000 nodes (or, 60 out of 1000 nodes) are seen to be isolated from the giant component across 100,000 experiments.*

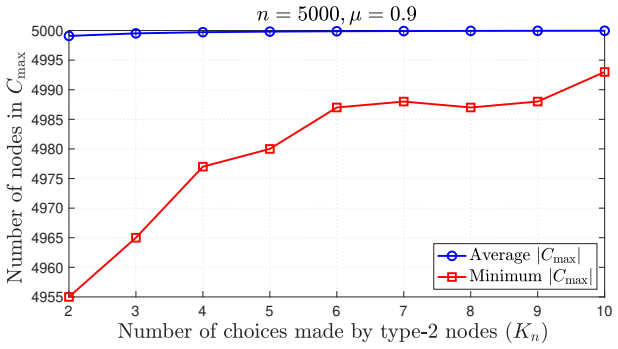

Fig. 3. *Average and minimum number of nodes contained in the largest connected component of $\mathbb{H}(n;\mu,K_n)$ across 100,000 experiments with $n=5000$, $\mu=0.9$ and $K_n\in\{2,\ldots,10\}$.*

IV. A Proof of Theorem 3.2

In this section, we provide a proof sketch for Theorem 3.2. Recall that $C_{\max}(n;\mu,K_n)$ denotes the largest connected component of $\mathbb{H}(n;\mu,K_n)$. We start by defining a *cut*.

Definition 4.1 (Cut): [28, Definition 6.3] Consider a graph $\mathcal{G}$ with the node set $\mathcal{N}$. A *cut* is defined as a non-empty subset $S\subset\mathcal{N}$ of nodes that is *isolated* from the rest

of the graph. Namely, $S \subset \mathcal{N}$ is a cut if there is no edge between S and $S^{\rm c} = \mathcal{N} \setminus S$.

It is clear from Definition 4.1 that if S is a cut, then so is $S^{\rm c}$. It is important to note the distinction between a *cut* as defined above and the notion of a *connected component* given in Definition 3.1. A connected component is isolated from the rest of the nodes by Definition 3.1 and therefore it is also a cut. However, nodes within a cut may not be connected meaning that not every cut is a connected component.

Let $\mathcal{E}_n(\mu, K_n; S)$ denote the event that $S \subset \mathcal{N}$ is a cut in $\mathbb{H}(n; \mu, K_n)$ as per Definition 4.1. Event $\mathcal{E}_n(\mu, K_n; S)$ occurs if no nodes in S pick neighbors in $S^{\rm c}$ and no nodes in S pick neighbors in $S^{\rm c}$. Thus, we have

$$\mathcal{E}_n(\mu, K_n; S) = \bigcap_{i \in S} \bigcap_{j \in S^c} \left(\{ i \notin \Gamma_{n,j} \} \cap \{ j \notin \Gamma_{n,i} \} \right).$$

Let $\mathcal{Z}(x_n; \mu, K_n)$ denote the event that $\mathbb{H}(n; \mu, K_n)$ has no cut $S \subset \mathcal{N}$ with size $x_n \leq |S| \leq n - x_n$ where $x : \mathbb{N}_0 \to \mathbb{N}_0$ is a sequence such that $x_n \leq n/2 \ \forall n$. In other words, $\mathcal{Z}(x_n; \mu, K_n)$ is the event that there are no cuts in $\mathbb{H}(n; \mu, K_n)$ whose size falls in the range $[x_n, n - x_n]$. Since if S is a cut, then so is S^c (i.e., if there is a cut of size m then there must be a cut of size $n - m$), we see that

$$\mathcal{Z}(x_n; \mu, K_n) = \bigcap_{S \in \mathcal{P}_n :\ x_n \leq |S| \leq \lfloor \frac{n}{2} \rfloor} \left(\mathcal{E}_n(\mu, K_n; S) \right)^{\rm c},$$

where $\mathcal{P}_n$ is the collection of all non-empty subsets of $\mathcal{N}$. Next, we present an upper bound on $\mathbb{P}\left[(\mathcal{Z}(M; \mu, K_n))^{\rm c} \right]$, i.e, the probability that there exists a cut with size in the range $[M, n - M]$ for $\mathbb{H}(n; \mu, K_n)$.

Proposition 4.2: Consider a scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$ such that $K_n \geq 2 \ \forall n$, $K_n = O(1)$, and $\mu \in (0, 1)$. It holds that

$$\begin{aligned} &\mathbb{P}\left[(\mathcal{Z}(M; \mu, K_n))^{\rm c} \right] \\ &\leq \frac{\exp\{ -M \left(\langle K_n \rangle - 1 \right) (1 - o(1)) \}}{1 - \exp\{ - \left(\langle K_n \rangle - 1 \right) (1 - o(1)) \}} + o(1). \end{aligned} \tag{9}$$

The proof of Proposition 4.2 is presented in Section V.

The following Lemma establishes the relevance of the event $\mathcal{Z}(x_n; \mu, K_n)$ in obtaining a lower bound for the size of the largest connected component.

Lemma 4.3: For any sequence $x : \mathbb{N}_0 \to \mathbb{N}_0$ such that $x_n \leq \lfloor n/3 \rfloor$ for all n, we have

$$\mathcal{Z}(x_n; \mu, K_n) \implies |C_{\max}(n; \mu, K_n)| > n - x_n.$$

Proof. Assume that $\mathcal{Z}(x_n; \mu, K_n)$ takes place, i.e., there is no cut in $\mathbb{H}(n; \mu, K_n)$ of size in the range $[x_n, n - x_n]$. Since a connected component is also a cut, this also means that there is no connected component of size in the range $[x_n, n - x_n]$. Since every graph has at least one connected component, it either holds that the largest one has size $|C_{\max}(n; \mu, K_n)| > n - x_n$, or that $|C_{\max}(n; \mu, K_n)| < x_n$. We now show that it must be the case that $|C_{\max}(n; \mu, K_n)| > n - x_n$ under the assumption that $x_n \leq n/3$. Assume towards a contradiction that $|C_{\max}(n; \mu, K_n)| < x_n$ meaning that the size of each connected component is less than x_n. Note that the union of any set of connected components is either a cut, or it spans the entire network. If no cut exists with size in the range $[x_n, n - x_n]$, then the union of any set of connected components should also have a size outside of $[x_n, n - x_n]$. Also, the union of all connected components has size n. Let $C_1, C_2, \ldots, C_{\max}$ denote the set of connected components in increasing size order. Let $m \geq 1$ be the largest integer such that $\sum_{i=1}^{m} |C_i| < x_n$. Since $|C_{m+1}| < x_n$, we have

$$x_n \leq \sum_{i=1}^{m+1} |C_i| < x_n + x_n \leq 2n/3 \leq n - x_n.$$

This means that $\cup_{i=1}^{m+1} C_i$ constitutes a cut with size in the range $[x_n, n - x_n]$ contradicting the event $\mathcal{Z}(x_n; \mu, K_n)$. We thus conclude that if $\mathcal{Z}(x_n; \mu, K_n)$ takes place with $x_n \leq n/3$, then we must have $|C_{\max}(n; \mu, K_n)| > n - x_n$. ■

We now have all the requisite ingredients for establishing Theorem 3.2. Substituting $x_n = M, \ \forall n$ in Lemma 4.3 for some finite integer M, we get $\mathcal{Z}(M; \mu, K_n) \implies |C_{\max}(n; \mu, K_n)| > n - M$. Equivalently, we have $|C_{\max}(n; \mu, K_n)| \leq n - M \implies \mathcal{Z}(M; \mu, K_n)^{\rm c}$. This gives

$$\mathbb{P}\left[|C_{\max}(n; \mu, K_n)| \leq n - M \right] \leq \mathbb{P}\left[\mathcal{Z}(M; \mu, K_n)^{\rm c} \right] \tag{10}$$

and we get (5) by using Proposition 4.2 in (10).

V. A Proof of Proposition 4.2

A. Useful Facts

For $0 \leq x < 1$ and for a sequence $y = 0, 1, 2 \ldots$, we have

$$1 - xy \leq (1 - x)^y \leq 1 - xy + \frac{1}{2} x^2 y^2. \tag{11}$$

A proof of (11) can be found in [39, Fact 2].

For all $x \in \mathbb{R}$, we have

$$1 \pm x \leq e^{\pm x}. \tag{12}$$

For $0 \leq m \leq n_1 \leq n_2, \quad m, n_1, n_2 \in \mathbb{N}_0$,

$$\frac{\binom{n_1}{m}}{\binom{n_2}{m}} = \prod_{i=0}^{m-1} \left(\frac{n_1 - i}{n_2 - i} \right) \leq \left(\frac{n_1}{n_2} \right)^m. \tag{13}$$

From [17, Fact 4.1], we have that for $r = 1, 2, \ldots, \lfloor \frac{n}{2} \rfloor$,

$$\binom{n}{r} \leq \left(\frac{n}{r} \right)^r \left(\frac{n}{n - r} \right)^{n - r}. \tag{14}$$

B. Proof of Proposition 4.2

Proof. Recall that we have

$$\mathcal{Z}(x_n; \mu, K_n) = \bigcap_{S \in \mathcal{P}_n :\ x_n \leq |S| \leq \lfloor \frac{n}{2} \rfloor} \left(\mathcal{E}_n(\mu, K_n; S) \right)^{\rm c},$$

where $\mathcal{Z}(x_n; \mu, K_n)$ denotes the event that $\mathbb{H}(n; \mu, K_n)$ has no cut $S \subset \mathcal{N}$ with size $x_n \leq |S| \leq n - x_n$. Taking the complement of both sides and using a union bound we get

$$\mathbb{P}\left[(\mathcal{Z}(x_n; \mu, K_n))^{\rm c} \right] \leq \sum_{S \in \mathcal{P}_n : x_n \leq |S| \leq \lfloor \frac{n}{2} \rfloor} \mathbb{P}[\mathcal{E}_n(\mu, K_n; S)]$$

$$= \sum_{r=x_n}^{\lfloor \frac{n}{2} \rfloor} \left(\sum_{S \in \mathcal{P}_{n,r}} \mathbb{P}[\mathcal{E}_n(\mu, K_n; S)] \right), \tag{15}$$

where $\mathcal{P}_{n,r}$ denotes the collection of all subsets of $\mathcal{N}$ with exactly r elements. For each $r = 1, \ldots, n$, we simplify the notation by writing $\mathcal{E}_{n,r}(\mu, K_n) = \mathcal{E}_n(\mu, K_n; \{1, \ldots, r\})$. From the exchangeability of the node labels and associated random variables, we get

$$\mathbb{P}[\mathcal{E}_n(\mu, K_n; S)] = \mathbb{P}[\mathcal{E}_{n,r}(\mu, K_n)], \quad S \in \mathcal{P}_{n,r}.$$

Noting that $|\mathcal{P}_{n,r}| = \binom{n}{r}$, we obtain

$$\sum_{S \in \mathcal{P}_{n,r}} \mathbb{P}[\mathcal{E}_n(\mu, K_n; S)] = \binom{n}{r} \mathbb{P}[\mathcal{E}_{n,r}(\mu, K_n)].$$

Substituting into (15) we obtain

$$\mathbb{P}\left[(\mathcal{Z}(x_n; \mu, K_n))^{\mathrm{c}}\right] \leq \sum_{r=x_n}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{r} \mathbb{P}[\mathcal{E}_{n,r}(\mu, K_n)]. \tag{16}$$

In view of (16), the proof for Proposition 4.2 will follow upon showing

$$\sum_{r=M}^{\lfloor n/2 \rfloor} \binom{n}{r} \mathbb{P}[\mathcal{E}_{n,r}(\mu, K_n)] \leq \frac{\exp\{-M\left(\langle K_n \rangle - 1\right)(1 - o(1))\}}{1 - \exp\{-\left(\langle K_n \rangle - 1\right)(1 - o(1))\}} + o(1). \tag{17}$$

We have

$$\begin{aligned}
&\binom{n}{r} \mathbb{P}[\mathcal{E}_{n,r}(\mu, K_n)] \\
&= \binom{n}{r} \left(\mu \left(\frac{n-r-1}{n-1} \right) + (1-\mu) \frac{\binom{n-r-1}{K_n}}{\binom{n-1}{K_n}} \right)^{n-r} \\
&\quad \cdot \left(\mu \left(\frac{r-1}{n-1} \right) + (1-\mu) \frac{\binom{r-1}{K_n}}{\binom{n-1}{K_n}} \right)^{r} \\
&\leq \binom{n}{r} \left(\mu \left(1 - \frac{r}{n-1} \right) + (1-\mu) \left(1 - \frac{r}{n-1} \right)^{K_n} \right)^{n-r} \\
&\quad \cdot \left(\mu \left(\frac{r-1}{n-1} \right) + (1-\mu) \left(\frac{r-1}{n-1} \right)^{K_n} \right)^{r} \qquad (18) \\
&\leq \binom{n}{r} \left(\mu \left(1 - \frac{r}{n} \right) + (1-\mu) \left(1 - \frac{r}{n} \right)^{K_n} \right)^{n-r} \\
&\quad \cdot \left(\mu \left(\frac{r}{n} \right) + (1-\mu) \left(\frac{r}{n} \right)^{K_n} \right)^{r} \\
&\leq \left(\frac{n}{r} \right)^{r} \left(\frac{n}{n-r} \right)^{n-r} \left(1 - \frac{r}{n} \right)^{n-r} \left(\frac{r}{n} \right)^{r} \\
&\quad \cdot \left(\mu + (1-\mu) \left(1 - \frac{r}{n} \right)^{K_n - 1} \right)^{n-r} \\
&\quad \cdot \left(\mu + (1-\mu) \left(\frac{r}{n} \right)^{K_n - 1} \right)^{r} \qquad (19) \\
&= \left(\mu + (1-\mu) \left(1 - \frac{r}{n} \right)^{K_n - 1} \right)^{n} \\
&\quad \cdot \left(\frac{\mu + (1-\mu) \left(\frac{r}{n} \right)^{K_n - 1}}{\left(\mu + (1-\mu) \left(1 - \frac{r}{n} \right)^{K_n - 1} \right)} \right)^{r} \\
&\leq \left(\mu + (1-\mu) \left(1 - \frac{r}{n} \right)^{K_n - 1} \right)^{n} \qquad (20)
\end{aligned}$$

where (18) uses (13), (19) follows from (14) and (20) is plain from the observation that $r/n \leq 1/2$.

We divide the summation in (17) into two parts depending on whether r exceeds $n/\log n$. The steps outlined below can be used to upper bound the summation in (17) for an arbitrary splitting of the summation indices.

$$\begin{aligned}
\sum_{r=M}^{\lfloor n/2 \rfloor} \binom{n}{r} \mathbb{P}[\mathcal{E}_{n,r}(\mu, K_n)] &= \sum_{r=M}^{\lfloor n/\log n \rfloor} \binom{n}{r} \mathbb{P}[\mathcal{E}_{n,r}(\mu, K_n)] \\
&\quad + \sum_{r=\lfloor n/\log n \rfloor}^{\lfloor n/2 \rfloor} \binom{n}{r} \mathbb{P}[\mathcal{E}_{n,r}(\mu, K_n)].
\end{aligned} \tag{21}$$

We first upper bound each term in the summation with indices in the range $M \leq r \leq \lfloor n/\log n \rfloor$.

Range 1: $M \leq r \leq \lfloor n/\log n \rfloor$

$$\begin{aligned}
&\binom{n}{r} \mathbb{P}[\mathcal{E}_{n,r}(\mu, K_n)] \\
&\leq \left(\mu + (1-\mu) \left(1 - \frac{r}{n} \right)^{K_n - 1} \right)^{n} \qquad (22) \\
&= \left(1 - (1-\mu) \left(1 - \left(1 - \frac{r}{n} \right)^{K_n - 1} \right) \right)^{n} \qquad (23)
\end{aligned}$$

For $r \leq \lfloor n/\log n \rfloor$, we have $\frac{r}{n} = o(1)$. Using Fact (11) with $x = \frac{r}{n}$ we get

$$\begin{aligned}
&\binom{n}{r} \mathbb{P}[\mathcal{E}_{n,r}(\mu, K_n)] \\
&\leq \left(1 - (1-\mu) \left(1 - \left(1 - \frac{r(K_n - 1)}{n} + \frac{r^2 (K_n - 1)^2}{2n^2} \right) \right) \right)^{n} \\
&= \left(1 - (1-\mu) \frac{r(K_n - 1)}{n} \left(1 - \frac{r(K_n - 1)}{2n} \right) \right)^{n}
\end{aligned}$$

Using $r \leq n/\log n$, (12) and that $K_n = O(1)$, we obtain,

$$\begin{aligned}
&\binom{n}{r} \mathbb{P}[\mathcal{E}_{n,r}(\mu, K_n)] \\
&\leq \left(1 - (1-\mu) \frac{r(K_n - 1)}{n} \left(1 - \frac{(K_n - 1)}{2 \log n} \right) \right)^{n} \\
&\leq \exp\left\{ -(1-\mu) r (K_n - 1) \left(1 - \frac{(K_n - 1)}{2 \log n} \right) \right\} \qquad (24) \\
&= \exp\left\{ -r(1-\mu)(K_n - 1)\left(1 - o(1)\right) \right\} \qquad (25) \\
&= \exp\left\{ -r(\langle K_n \rangle - 1)\left(1 - o(1)\right) \right\}. \qquad (26)
\end{aligned}$$

Next, we upper bound the second term in the summation (21) with indices in the range $\lfloor n/\log n \rfloor + 1 \leq r \leq \lfloor n/2 \rfloor$.

Range 2: $\lfloor n/\log n \rfloor + 1 \leq r \leq \lfloor n/2 \rfloor$

Observe that

$$\begin{aligned}
\binom{n}{r}\mathbb{P}[\mathcal{E}_{n,r}(\mu, K_n)] &\leq \left(\mu + (1-\mu)\left(1 - \frac{r}{n}\right)^{K_n - 1}\right)^n \\
&\leq \left(\mu + (1-\mu)\left(1 - \frac{r}{n}\right)\right)^n \qquad (27) \\
&= \left(1 - \frac{r}{n}(1-\mu)\right)^n \\
&\leq \exp\left(-r(1-\mu)\right) \qquad (28) \\
&= o(1), \qquad (29)
\end{aligned}$$

where (28) follows from noting that $K_n \geq 2$ and (27) is a consequence of (12). Finally, we use (26) and (29) in (21) as follows.

$$\begin{aligned}
&\sum_{r=M}^{\lfloor n/2 \rfloor} \binom{n}{r}\mathbb{P}[\mathcal{E}_{n,r}(\mu, K_n)] \\
&= \sum_{r=M}^{\lfloor n/\log n \rfloor} \binom{n}{r}\mathbb{P}[\mathcal{E}_{n,r}(\mu, K_n)] \\
&+ \sum_{r=\lfloor n/\log n \rfloor}^{\lfloor n/2 \rfloor} \binom{n}{r}\mathbb{P}[\mathcal{E}_{n,r}(\mu, K_n)] \\
&\leq \sum_{r=M}^{\lfloor n/\log n \rfloor} \exp\left\{-r(\langle K_n \rangle - 1)\left(1 - o(1)\right)\right\} + \sum_{r=\lfloor n/\log n \rfloor}^{\lfloor n/2 \rfloor} o(1) \\
&= \left(\sum_{r=M}^{\lfloor n/\log n \rfloor} \exp\left\{-r(\langle K_n \rangle - 1)\left(1 - o(1)\right)\right\}\right) + o(1) \\
&\leq \left(\sum_{r=M}^{\infty} \exp\left\{-r(\langle K_n \rangle - 1)\left(1 - o(1)\right)\right\}\right) + o(1).
\end{aligned}$$

Observe that the above geometric series has each term strictly less than one, and thus it is summable. This gives

$$\begin{aligned}
&\binom{n}{r}\mathbb{P}[\mathcal{E}_{n,r}(\mu, K_n)] \\
&\leq \frac{\exp\{-M\left(\langle K_n \rangle - 1\right)\left(1 - o(1)\right)\}}{1 - \exp\{-\left(\langle K_n \rangle - 1\right)\left(1 - o(1)\right)\}} + o(1).
\end{aligned}$$

■

VI. Conclusions

This work analyzes the existence and size of the giant component for the inhomogeneous random K-out graph. In particular, we prove that whenever $K_n \geq 2$, the largest connected sub-network spans all but finitely many nodes of the network with high probability. This result complements the existing results on the connectivity of inhomogeneous random K-out graphs. An open direction is characterizing the asymptotic size of the largest connected component of homogeneous K-out random graph when $K = 1$ which is not known to the best of our knowledge. For the inhomogeneous random K-out graph with r classes, an explicit bound on the probability of the number of nodes outside the giant component exceeding M would also be of interest.

Acknowledgments

We thank Prof. Giulia Fanti for insightful discussions on payment channel networks. This work has been supported in part by the National Science Foundation through grant CCF #1617934.

References

[1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex networks: Structure and dynamics," *Physics reports*, vol. 424, no. 4-5, pp. 175–308, 2006.

[2] A. Goldenberg, A. X. Zheng, S. E. Fienberg, E. M. Airoldi *et al.*, "A survey of statistical network models," *Foundations and Trends in Machine Learning*, vol. 2, no. 2, pp. 129–233, 2010.

[3] M. E. Newman, D. J. Watts, and S. H. Strogatz, "Random graph models of social networks," *Proceedings of the National Academy of Sciences*, vol. 99, no. suppl 1, pp. 2566–2572, 2002.

[4] S. M. Kakade, M. Kearns, L. E. Ortiz, R. Pemantle, and S. Suri, "Economic properties of social networks," in *Advances in Neural Information Processing Systems*, 2005, pp. 633–640.

[5] 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*, ser. CCS '02. New York, NY, USA: ACM, 2002, pp. 41–47. [Online]. Available: http://doi.acm.org/10.1145/586110.586117

[6] O. Yağan, "Random graph modeling of key distribution schemes in wireless sensor networks," Ph.D. dissertation, 2011.

[7] Y. Wang, G. Attebury, and B. Ramamurthy, "A survey of security issues in wireless sensor networks," *IEEE Communications Surveys Tutorials*, vol. 8, no. 2, pp. 2–23, Second 2006.

[8] Y. Xiao, V. K. Rayi, B. Sun, X. Du, F. Hu, and M. Galloway, "A survey of key management schemes in wireless sensor networks," *Computer Communications*, vol. 30, pp. 2314 – 2341, 2007, special issue on security on wireless ad hoc and sensor networks. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0140366407001752

[9] T. I. Fenner and A. M. Frieze, "On the connectivity of random m-orientable graphs and digraphs," *Combinatorica*, vol. 2, no. 4, pp. 347–359, Dec 1982.

[10] B. Bollobás, *Random graphs*. Cambridge university press, 2001, vol. 73.

[11] O. Yağan and A. M. Makowski, "On the connectivity of sensor networks under random pairwise key predistribution," *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5754–5762, Sept 2013.

[12] ——, "Modeling the pairwise key predistribution scheme in the presence of unreliable links," *Information Theory, IEEE Transactions on*, vol. 59, no. 3, pp. 1740–1760, March 2013.

[13] F. Yavuz, J. Zhao, O. Yağan, and V. Gligor, "k-connectivity in random k-out graphs intersecting erdős-rényi graphs," *IEEE Transactions on Information Theory*, vol. 63, no. 3, pp. 1677–1692, 2017.

[14] ——, "Toward k-connectivity of the random graph induced by a pairwise key predistribution scheme with unreliable links," *IEEE Transactions on Information Theory*, vol. 61, no. 11, pp. 6251–6271, 2015.

[15] H. Chan, A. Perrig, and D. Song, "Random key predistribution schemes for sensor networks," in *Proc. of IEEE S&P 2003*, 2003.

[16] G. Fanti, S. B. Venkatakrishnan, S. Bakshi, B. Denby, S. Bhargava, A. Miller, and P. Viswanath, "Dandelion++: Lightweight cryptocurrency networking with formal anonymity guarantees," *Proc. ACM Meas. Anal. Comput. Syst.*, vol. 2, no. 2, pp. 29:1–29:35, Jun. 2018.

[17] R. Eletreby and O. Yağan, "Connectivity of wireless sensor networks secured by the heterogeneous random pairwise key predistribution scheme," in *Proc. of IEEE CDC 2018*, Dec 2018.

[18] R. Eletreby and O. Yağan, "On the connectivity of inhomogeneous random k-out graphs," in *2019 IEEE International Symposium on Information Theory (ISIT)*, July 2019, pp. 1482–1486.

[19] X. Du, Y. Xiao, M. Guizani, and H.-H. Chen, "An effective key management scheme for heterogeneous sensor networks," *Ad Hoc Networks*, vol. 5, no. 1, pp. 24–34, 2007.

[20] R. Eletreby and O. Yağan, "k-connectivity of inhomogeneous random key graphs with unreliable links," *IEEE Transactions on Information Theory*, vol. 65, no. 6, pp. 3922–3949, June 2019.

[21] ——, "Connectivity of wireless sensor networks secured by heterogeneous key predistribution under an on/off channel model," *IEEE Transactions on Control of Network Systems*, 2018.

[22] O. Yağan, "Zero-one laws for connectivity in inhomogeneous random key graphs," *IEEE Transactions on Information Theory*, vol. 62, no. 8, pp. 4559–4574, Aug 2016.
[23] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *science*, vol. 286, no. 5439, pp. 509–512, 1999.
[24] K. Lu, Y. Qian, M. Guizani, and H.-H. Chen, "A framework for a distributed key management scheme in heterogeneous wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 2, pp. 639–647, February 2008.
[25] 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*, 2007, pp. 78–88.
[26] M. Yarvis, N. Kushalnagar, H. Singh, A. Rangarajan, Y. Liu, and S. Singh, "Exploiting heterogeneity in sensor networks," in *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, vol. 2, March 2005, pp. 878–890 vol. 2.
[27] J. Hwang and Y. Kim, "Revisiting random key pre-distribution schemes for wireless sensor networks," in *Proceedings of the 2nd ACM workshop on Security of ad hoc and sensor networks*. ACM, 2004, pp. 43–52.
[28] A. Mei, A. Panconesi, and J. Radhakrishnan, "Unassailable sensor networks," in *Proc. of the 4th International Conference on Security and Privacy in Communication Netowrks*, ser. SecureComm '08. New York, NY, USA: ACM, 2008.
[29] X. Liu, "Coverage with connectivity in wireless sensor networks," in *2006 3rd International Conference on Broadband Communications, Networks and Systems*. IEEE, 2006, pp. 1–8.
[30] P. Erdős and A. Rényi, "On the evolution of random graphs," *Publ. Math. Inst. Hung. Acad. Sci*, vol. 5, no. 1, pp. 17–60, 1960.
[31] J. Poon and T. Dryja, "The bitcoin lightning network: Scalable off-chain instant payments," 2016.
[32] I. A. Seres, L. Gulyás, D. A. Nagy, and P. Burcsi, "Topological analysis of bitcoin's lightning network," *CoRR*, vol. abs/1901.04972, 2019. [Online]. Available: http://arxiv.org/abs/1901.04972
[33] W. Tang, W. Wang, G. Fanti, and S. Oh, "Privacy-utility tradeoffs in routing cryptocurrency over payment channel networks," *arXiv preprint arXiv:1909.02717*, 2019.
[34] V. Sivaraman, S. B. Venkatakrishnan, K. Ruan, P. Negi, L. Yang, R. Mittal, M. Alizadeh, and G. Fanti, "High throughput cryptocurrency routing in payment channel networks," 2018.
[35] J. A. Bondy, U. S. R. Murty *et al.*, *Graph theory with applications*. Macmillan London, 1976, vol. 290.
[36] R. Van Der Hofstad, *Random graphs and complex networks*. Cambridge university press, 2016, vol. 1.
[37] S. Janson, "Probability asymptotics: notes on notation," *arXiv preprint arXiv:1108.3924*, 2011.
[38] K. Rybarczyk, "Diameter, connectivity, and phase transition of the uniform random intersection graph," *Discrete Mathematics*, vol. 311, no. 17, pp. 1998–2019, 2011.
[39] J. Zhao, O. Yağan, and V. Gligor, "k-connectivity in random key graphs with unreliable links," *IEEE Transactions on Information Theory*, vol. 61, no. 7, pp. 3810–3836, July 2015.
[40] K. Rybarczyk, "Sharp threshold functions for random intersection graphs via a coupling method," *the electronic journal of combinatorics*, vol. 18, no. 1, p. 36, 2011.

APPENDIX

A. Mean node degree in $\mathbb{H}(n;\mu,K_n)$

The probability that node i picks node j where $i,j\in\mathcal{N}$ depends on the type of node i and is given by

$$\mathbb{P}[j\in\Gamma_{n,i}]=\mu\frac{1}{n-1}+(1-\mu)\frac{K_n}{n-1}=\frac{\langle K_n\rangle}{n-1}. \quad (30)$$

Let $i\sim j$ denote the event that node i can securely communicate with node j.

$$\begin{aligned}\mathbb{P}[i\sim j]&=1-(1-\mathbb{P}[i\in\Gamma_{n,j}])(1-\mathbb{P}[j\in\Gamma_{n,i}]),\\&=1-\left(1-\frac{\langle K_n\rangle}{n-1}\right)^2,\\&=\frac{2\langle K_n\rangle}{n-1}-\left(\frac{\langle K_n\rangle}{n-1}\right)^2. \quad (31)\end{aligned}$$

Consequently, the mean degree of node i, when $K_n=o(1)$ is $(1-o(1))2\langle K_n\rangle$.

B. Inhomogeneous random K-out graph with r classes

Here, each node belongs to type-i independently with probability μ_i for $i=1,\ldots,r$ and $\sum_{i=1}^r\mu_i=1$. Each type-i nodes gets paired with $K_{i,n}$ other nodes, chosen uniformly at random from among all other nodes where $1\le K_{1,n}<K_{2,n}<\ldots<K_{r,n}$. Let $\boldsymbol{K_n}$ denote $[K_{1,n},K_{2,n},\ldots,K_{r,n}]$ and $\boldsymbol{\mu}=[\mu_1,\mu_2,\ldots,\mu_r]$ with $\mu_i>0$.

Corollary 6.1: If $K_{r,n}\ge 2\,\forall n$ then for the inhomogeneous random K-out graph $\mathbb{H}(n;\boldsymbol{\mu},\boldsymbol{K}_n)$ with r node types, we have

$$|C_{\max}(n;\boldsymbol{\mu},\boldsymbol{K}_n)|=n-O(1)\ \text{ whp.}$$

Proof. The proof involves showing the existence of a coupling between the graphs $\mathbb{H}(n;\mu,K_n)$ and $\mathbb{H}(n;\boldsymbol{\mu},\boldsymbol{K}_n)$ such that the edge set of $\mathbb{H}(n;\mu,K_n)$ is contained in the edge set of $\mathbb{H}(n;\boldsymbol{\mu},\boldsymbol{K}_n)$. For any monotone-increasing property $\mathcal{P}$, i.e., a property which holds upon addition of edges to the graph (see [40, p. 13]) we have

$$\begin{aligned}&\mathbb{P}[\mathbb{H}(n;\boldsymbol{\mu},\boldsymbol{K}_n)\text{ has property }\mathcal{P}]\\&\ge\mathbb{P}[\mathbb{H}(n;\mu,K_n)\text{ has property }\mathcal{P}] \quad (32)\end{aligned}$$

It is plain that the property $|C_{\max}(n;\mu,K_n)|\ge n-M$ is monotone increasing upon edge addition. Therefore, if there exists a *coupling* under which $\mathbb{H}(n;\mu,K_n)$ is a spanning subgraph of $\mathbb{H}(n;\boldsymbol{\mu},\boldsymbol{K}_n)$; i.e., if we can generate an instantiation of $\mathbb{H}(n;\boldsymbol{\mu},\boldsymbol{K}_n)$ by adding edges to an instantiation of $\mathbb{H}(n;\mu,K_n)$, then we can use (32) to establish this Corollary. Let $\tilde{\mu}$ denote $\sum_{i=1}^{r-1}\mu_i$. Consider an instantiation of an inhomogeneous random graph $\mathbb{H}(n;\tilde{\mu},K_{r,n})$ with two classes such that each of the n nodes is independently assigned as type-1 (resp., type-2) with probability $\tilde{\mu}$ (resp., $1-\tilde{\mu}$) and then type-1 (resp., type-2) nodes draw edges to 1 (resp. $K_{r,n}$) nodes chosen uniformly at random. From this instantiation, we can generate an instantiation of $\mathbb{H}(n;\boldsymbol{\mu},\boldsymbol{K}_n)$ as follows. First, let each type-1 node be independently reassigned as type-i with probability $\frac{\mu_i}{\tilde{\mu}}$ for $i=1,2,\ldots,r-1$. Next, for $i=2,\ldots,r-1$, let each type-i node pick $K_{i,n}-1$ additional neighbors that were not chosen by it initially. Next, we draw an undirected edge between each pair of nodes where at least one picked the other. Clearly, this process creates a graph whose edge set is a superset of the edge set of the realization of $\mathbb{H}(n;\tilde{\mu},K_{r,n})$ that we started with. In addition, in the new graph, the probability of a node picking $K_{i,n}$ other nodes (i.e., being type-i) is given by $\tilde{\mu}\frac{\mu_i}{\tilde{\mu}}=\mu_i$, for $i=1,2,\ldots,n$. We thus conclude that the new graph obtained constitutes a realization of $\mathbb{H}(n;\boldsymbol{\mu},\boldsymbol{K}_n)$. Since, the initial realization of $\mathbb{H}(n;\tilde{\mu},K_{r,n})$ was arbitrary, this establishes the desired coupling argument and we conclude that (32) holds for the property $|C_{\max}(n;\mu,K_n)|\ge n-M$. ■