DELIS
Dynamically Evolving, Large-scale Information Systems

Integrated Project
Member of the FET Proactive Initiative Complex Systems

DELIS-TR-362

Unassailable Sensor Networks

A. Mei, A. Panconesi, J. Radhakrishnan

2006
Unassailable Sensor Networks

ALESSANDRO MEI, ALESSANDRO PANCONESI
University of Rome “La Sapienza”, Italy
JAIKUMAR RADHAKRISHNAN
Tata Institute of Fundamental Research, Bombay, India

Abstract—In this paper we prove in a mathematically rigorous way that random pre-distribution of keys have very strong security properties against massive attacks by the so called omniscient adversary. The omniscient adversary knows the key ring of every sensor of the network and operates its choices during the attack in the optimal (most disruptive) way. We consider two important security properties: We say that the network is unassailable if the adversary cannot compromise a linear fraction of the communication links by compromising a sub-linear fraction of the nodes, and that the network is unsplitable if the adversary cannot partition the network into two (or more) linear size fragments. We show how to set the relevant parameters—pool and key ring size—in such a way that the network is not only connected, but also provably unassailable and unsplitable with high probability against the omniscient adversary, simultaneously. Moreover, we also show how to set the parameters in such a way to form a giant component in the network, a connected subgraph including, say, 99% of the sensors. Giant components in kryptographs are sparse, emerge by using much smaller key rings, and, quite remarkably, are provably unassailable and unsplitable as well. All these results are supported by experiments.

Index Terms—Wireless security, sensor networks.

I. INTRODUCTION

Key management is one of the central security issues for wireless sensor networks (see [17], among others). The so-called random pre-distribution of keys introduced in [10] is one of the most promising approaches that have been proposed so far to set up pairwise keys in a sensor network. This approach seems to be general enough to be applicable to other contexts too, such as peer-to-peer networks, and therefore deserves to be studied. In this paper we prove in a mathematically rigorous fashion that the model has very strong security and fault-tolerant properties and present experiments to back up our main findings. To explain our results we review briefly the model, formally defined in Section III. We are given a network of \( n \) nodes and a pool of \( K \) cryptographic keys. Each node \( u \) is given a random subset \( K_u \) of \( k \) keys, its key ring. Two nodes in the network are connected by a link if and only if they are within transmission radius and share a key. The resulting graph is called a kryptograph. The basic question is how to set the values of \( K \) and \( k \) in order to have some desirable properties, such as various forms of connectivity and, simultaneously, security. It should be noted that security and connectivity are at odds. The former likes large key rings (relative to the pool size), since this increases the probability of having links. On the other hand, the latter likes them to be small, since in this fashion capturing one sensor is unlikely to corrupt many links. As shown in [6], if

\[
\frac{k^2}{K} \sim \frac{\log n}{n}
\]  

then the network is connected with high probability, while it is likely to be disconnected below this threshold.

In this paper we consider an extremely severe kind of (massive) attack. The adversary selects a certain number of nodes with the aim of compromising links. Key rings of the selected nodes belong to the adversary. We assume pessimistically that a link \( uv \) is compromised as soon as the adversary owns a key in \( K_u \cap K_v \). The assumption is pessimistic because this key may or may not be used by \( u \) and \( v \). We also assume that the adversary knows the sets \( K_u \), i.e. it knows how the keys are distributed. Its aim is to corrupt linearly many links. The question is: How many nodes does it have to capture?

We call this the omniscient adversary. In this paper we prove in a mathematically rigorous way that in order to compromise a large (linear fraction) of the links the omniscient adversary is forced with overwhelming probability to capture a large (linear fraction) of the nodes, hence the term unassailable. This holds assuming Condition 1. Therefore kryptographs are at the same time connected and secure (in the above sense) with high probability. Our proofs show that, in essence, even if the adversary knows the composition of the key rings and operates its choices in the optimal (most disruptive) way, capturing a node essentially corrupts only the links incident on the captured node.

It might be objected that this adversary is unrealistically strong. This however is entirely desirable! If we are able to certify the security of the network against such a strong adversary it follows that it will be secure against more realistic attacks. For instance the result holds for an adversary that captures nodes blindly, at random. In fact, it implies that to compromise linearly many links, linearly many nodes are needed, no matter how the adversary selects the nodes. Also, assuming such a strong adversary has important practical consequences. In the literature, much attention has been given to the problem of finding protocols for key discovery between neighbors in such a way that no information is released to the adversary. One example is challenge response [10],
that is energy consuming ($k$ messages and $k^2$ de-
cryptions per peer to discover shared keys) but keeps
key distribution as secret as possible. Our results on the
omniscient adversary show that, if we are interested in
protecting the network against massive attacks like the
one described above, we do not (asymptotically) gain any
advantage by using challenge response with respect to
any light-weight, non-cryptographic protocol that does
not keep key distribution secret. Broadcasting lists of
key indices in the clear is good as well (within constant
factors).

We prove that cryptographs enjoy another strong
security property. Suppose now that the aim of the
adversary is to split the network into two large sets
and to compromise all links between them, i.e. the
adversary wants to partition the network. We show that,
with overwhelming probability, the omniscient adversary
cannot do this. This result is a consequence of the
following structural property of cryptographs: with high
probability and simultaneously for all large (linear size)
sets of vertices, the number of edges between a set and its
complement is $O(n \log n)$. This fact implies good fault-
tolerant properties: to partition a network into two large
sets, a huge (linear fraction) of the links must go down.

For many applications it suffices to have a giant
connected component that covers the area within which
the network is deployed. We show that if
\[
\frac{k^2}{K} \sim \frac{1}{n}
\] (2)
then, with high probability, the network has a giant
connected component, say, a connected component con-
taining 99% of the vertices. Note that Condition 2 can be
satisfied with much smaller key rings than Condition 1.
This results in memory saving (smaller key ring) and
computation saving (faster key discovery), an important
consideration in resource starving environments typical
of sensor networks. But more importantly, networks
generated with Condition 2 are much sparser, not only
globally but also locally at the neighborhood level, than
those set up under Condition 1. This is very beneficial
because it translates in a reduced amount of interference,
limiting the number of packet collisions and correspond-
ing retransmissions. Nonetheless, we prove that the giant
connected component that emerge with high probability
by setting the parameters according to Condition 2 are
unassailable and unsplitable against the omniscient
adversary as well, just like the whole cryptograph.

The property we have been describing hold in the
general case, but they are formally proven for the full-
visibility case only. We made this choice because in this
fashion the basic argument with the underlying reasons
emerge more clearly, without cluttering technicalities.
The general case is dealt with in the experiments.

II. RELATED WORK

The idea of probabilistic key sharing for WSNs is
introduced by Eschenauer and Gligor [10]. The authors
also provide a simple and centralized algorithm for re-
keying in a distributed WSN. Later, in [4], a few new
mechanisms are described in the framework of random
key pre-distribution. Among these, the $q$-composite ran-
dom key pre-distribution scheme, a modification of the
basic scheme in [10], achieves better security under small
scale attacks while trading off increased vulnerability in
the face of a large scale physical attack on the network
sensors. Secure key discovery protocols for random pre-
distribution of keys have been proposed in [10], [4],
and [7].

Two schemes combine the random pre-distribution
scheme with a deterministic technique to build up secure
pairwise channels. The first scheme is proposed in [8].
The authors use a deterministic protocol proposed by
Blom [2] that allows any pair of nodes in a network to
find a pairwise secret key. As a salient feature, Blom’s
scheme guarantees a so called $\lambda$-secure property: as long
as no more than $\lambda$ nodes are compromised, the network
is perfectly secure. A $\lambda$-secure data structure built this
way is called a key space. The authors in [8] create a set
$\mathcal{W}$ composed of $\omega$ key spaces, and randomly assign up
to $\tau$ spaces per sensor. Two nodes can find a common
secret key if they have picked a common key space.
The second scheme is proposed in [15]. In principle, this
work is similar to [8], where Blundo et al’s polynomial
scheme [3] is used instead of Blom’s.

The problem of network connectivity when using ran-
dom pre-distribution of keys have been addressed in [10].
Their basic idea is that the network can be considered to
be a random graph in the sense of Erdős and Rényi [9].
However, notice that a cryptograph is generated by a
completely different random process and it is not clear
that this process can be approximated and if so, to what
extent, by a random graph in the sense of Erdős and
Rényi. In fact, random graphs and cryptographs have dif-
ferent structural properties as illustrated in [6]. There are
other difficulties. The Erdős-Rényi model assumes full-
visibility– any two devices can be connected by a direct
link regardless of their geographical position. There is no
guarantee that the Erdős-Renyi Theorem as used in [10]
ensures high probability of connectivity in the general
case, when devices are not within transmission range.
Later, the work in [12], that applies the well-known
results by Erdős and Rényi on giants components in
the random graph to the cryptograph following the same
methodology in [10], shows exactly the same problems.
We remark that having a precise understanding of a
model is always important, but especially so when secu-
rit y is at stake. Moreover, note that security properties,
like the ones we consider in this paper, cannot even be
formulated in the Erdős-Renyi model.

In [6], for the first time a precise mathematical anal-
alysis is given of the connectivity and security properties
of sensor networks that make use of the random pre-
distribution of keys. In that work, the authors show that,
then, with high probability, the network is simultaneously connected and secure against the random adversary, an adversary that blindly picks sensors at random to compromise the network.

Connectivity properties have been studied for non-secure wireless sensor networks as well. In [1], a geometric random model has been used to investigate minimum node degree and $h$-connectivity. Using a recent asymptotic result from Penrose [16], Bettstetter experimentally shows how to compute a communication range for a given number of nodes and a given integer $h$, the network is guaranteed to be $h$-connected. Equivalently, it is possible to compute how many sensors are needed to cover a given geographical area with an $h$-connected network.

In 1945, E Marczewski (see [14]) considered graphs where sets associated with vertices and two vertices were connected if their associated sets had an element in common. Recently, graphs obtained by choosing the sets randomly have been investigated [14], [13], [18], [19]. In these works, the sets associated with the vertices are usually large. For a certain choice of parameters this model of random graphs is shown to be similar to the $G(n,p)$ model of Erdős-Rényi. However, for the range of parameters of interest to us these results are not applicable.

### III. Preliminaries

In this paper we will study probability events that depend on $n$, the number of vertices of the network. Examples of such events are “the network is connected”, “the network has a giant component” etc.

**Definition 3.1:** Let $\mathcal{E}_n$ denote an event that depends on $n$. We will say that $\mathcal{E}_n$ holds with high probability if $\lim_{n\to\infty} \Pr[\mathcal{E}_n] = 1$.

The following definition captures exactly the kind of networks that are generated with random pre-distribution of keys.

**Definition 3.2:** Let $K$ be the size of a finite set of keys (the pool), and let $k \leq K$ be a fixed parameter. Let $[K] = \{1, 2, \ldots, K\}$ be the index set of the keys in the common pool of size $K$. The graph $G_{r,k,K}$ is defined as the geometric random graph obtained by the following procedure:

- First, each node $u$ is assigned a subset of keys, its key ring $K_u$, which is generated by sampling $[K]$ without replacement $k$ times.
- Second, the $n$ nodes are distributed uniformly at random on the given square geographical area, that, without loss of generality, we assume to be of side one (called the unit square).
- Third, $uv$ is an edge if (a) the two nodes are within distance $r$; and (b) $K_u \cap K_v \neq \emptyset$.

The resulting graph $G_{r,k,K}$ is called a kryptograph with parameters $r$, $k$, $K$ and $n$. In the special case in which every two nodes are within transmission range $r \geq \sqrt{2}$, the so-called full visibility case, the resulting graph is denoted as $G_r^k$.

In the sequel, for the sake of simplicity we shall identify $[K]$ with the set of keys and $K_u$ with the key ring of a vertex $u$. Note that all links of $G_{r,k,K}$ are secure by definition (edge $uv$ exists only if vertices $u$ and $v$ share at least one key). Therefore if the kryptograph is connected it is so via secure links alone.

In this paper we study analytically only the full visibility case, even though our proofs extend to the general case. We will study the network under the assumptions

$$\frac{k^2}{K} \sim \frac{\log n}{n}$$

and

$$\frac{k^2}{K} \sim \frac{1}{n}.$$  

The first is a necessary and sufficient condition if we want full connectivity with high probability [6]. We show in this paper that the second condition ensures that, with high probability, the network has a giant component. We will also assume that $k \geq 2$.

**Definition 3.3:** The omniscient adversary is defined as follows:

- It knows the key ring of every vertex in the network;
- It selects a set of $t$ vertices, thereby capturing all their key rings;
- A link $uv$ is corrupted if the adversary has captured a key $x \in K_u \cap K_v$ (one key is enough);
- The adversary must corrupt a constant fraction of the links: Its goal is to minimize $t$.

Henceforth we will refer to the omniscient adversary simply as the adversary.

**Definition 3.4:** A network is unassailable if the adversary must collect a linear fraction of the nodes in order to corrupt a linear fraction of the links.

We will investigate another property that is useful not only for security but also for fault-tolerance. Assume that the aim of the adversary now is to find a set of vertices $S$ such that all edges from $S$ to $V - S$ are corrupted. Then we say that the set $S$ is bad.

**Definition 3.5:** A network is unsplittable if there is no bad set $S$ of size $|S| \leq \frac{1}{10}$, where $\beta := \frac{1}{10}$.

We will show that if the adversary captures $t = o(n)$ many vertices then, with high probability the network is unsplittable. We remark that this result holds for all $\beta \in (0, \frac{1}{2})$. Our choice of $\beta = \frac{1}{10}$ is for the sake of simplicity.

We now review some standard definitions and facts.

**Definition 3.6:** We use $f \sim g$ to mean that $f$ and $g$ are the same up to constant factors. We say that $f(n) = o(g(n))$ if $\lim_{n \to \infty} \frac{f(n)}{g(n)} = 0$.

**Fact 3.7 (Hoeffding bound [11], [5]):** Let $K_1$, $K_2$, $K_3$, $K_4$, $K_5$, $K_6$ be subsets of $[K]$, each of size $pK$ (for some
\( p \in [0, 1] \), chosen independently at random. Let \( A \) be a subset of \([K]\) of size \( a \). Then, for all \( t \in [0, 1] \),
\[
\Pr \left[ \sum_{i=1}^{\ell} (K_i \cap A) \geq (p+t)la \right] \leq \left( \frac{ep}{p+t} \right)^{(p+t)la}.
\]
\[
\Pr \left[ \sum_{i=1}^{\ell} (K_i \cap A) \leq (1-t)pla \right] \leq \exp \left( -\frac{t^2}{2}pla \right).
\]

**Fact 3.8:**
\[
\binom{n}{k} \leq \left( \frac{en}{k} \right)^k.
\]

**IV. UNASSAILABLE NETWORKS**

A kryptograph can be thought of as a balls-and-bins experiment in which bins correspond to keys of the pool. To choose its key-ring, each vertex is given \( k \) balls. The key-ring is chosen by throwing the balls of a given vertex into \( k \) distinct bins. If a ball of vertex \( u \) lands in bin \( i \) then key \( i \) is assigned to \( u \). Thus, there is an edge \( uv \) in the graph iff there is a bin that contains a ball from \( u \) and a ball from \( v \).

The following observation is the key to limiting the power of the adversary. Let \( A \) denote the set of \( t \) bins (corresponding to the \( t \) keys) chosen by the adversary and let \( L \) denote the \( t \) largest bins. Then, denoting with \( B_i \) the number of balls in bin \( i \), we have
\[
\text{(#edges captured)} \leq \sum_{i \in A} \left( \frac{B_i}{2} \right) \leq \sum_{i \in L} \left( \frac{B_i}{2} \right).
\]

Thus, in order to show that the adversary with \( t \) keys cannot capture many edges, it is sufficient to show that the number of pairs in the \( t \) largest bins is small.

**Definition 4.1:** A bin is **useful** if it contains at least two balls. It is **useless** otherwise.

Our argument to show that the kryptograph is unassailable has two two parts. We show that: (a) in order to capture a linear fraction of the edges, the adversary must pick \( \Omega(n \log n) \) bins; (b) no vertex can contribute more than \( O(\log n) \) useful bins. It follows that in order to capture a linear fraction of the edges, a linear fraction of the vertices must be captured. Both (a) and (b) will be shown to hold with high probability. We begin with (b).

**Lemma 4.2:** With high probability, no vertex contributes more than \( 10 \log n \) useful bins.

**Proof:** For a vertex \( u \), let \( K_u \) denote the set of bins where the balls originating from \( u \) fall. Now fix some vertex \( u \), and let \( Y \) denote the number of useful bins in \( K_u \). Let \( X \) be the number of balls originating from the other vertices that land in one of the bins in \( K_u \). Clearly \( X \geq Y \), so that \( \Pr[Y > t] \leq \Pr[X > t] \). We will bound the latter. Note that
\[
X = \sum_{v : v \neq u} (K_v \cap K_u) \geq |K_v| \cap K_u|.
\]

Using Fact 3.7 (with \( p = \frac{k}{n} \), \( \ell = n - 1 \) and \( p + t = \frac{10 \log n}{(n-1)k} \geq 10p \)), we conclude that
\[
\Pr[X > 10 \log n] \leq \left( \frac{e}{10} \right)^{10 \log n} \ll n^{-10}.
\]

The claim follows from this by using the union bound, by summing over all choices for \( u \).

Note that the above lemma says that a huge fraction of the key ring is not used. For instance, if \( K \sim n^3 \) then \( k \sim n \sqrt{\log n} \), while only \( \log n \) keys from any one key ring are used to set up secure links.

Now we turn to (a). We have two claims: the first shows that the kryptographs has \( \Omega(n \log n) \) edges; the second shows that to cover \( \Omega(n \log n) \) edges, the adversary has to pick \( \Omega(n \log n) \) bins, because any one bin contributes only a constant number of edges.

**Lemma 4.3:** With high probability, a kryptograph has \( \Omega(n \log n) \) edges.

**Proof:** See Lemma 5.1 below.

It remains to show that any one bin cannot contribute more than a constant number of edges. The argument is simple if we assume that \( K \geq n^2 \). The claim holds even if \( n \log n < K \), but we do not include the more involved proof for this case for the sake of brevity.

**Lemma 4.4:** Assume \( K := n^\alpha \) with \( \alpha \geq 2 \). Then, with high probability, no bin contains more than \( 5 \) balls.

**Proof:** Let \( \mathcal{E}(i) \) denote the event that some bin contains at least \( t \) balls. Then, using the union bound we have
\[
\Pr[\mathcal{E}(i)] \leq K \Pr[\text{a bin has at least } t \text{ balls}]
\]
\[
\leq K \binom{n}{t} \left( \frac{k}{K} \right)^t
\]
\[
= K \left( \frac{en}{t} \right)^t \left( \frac{k}{K} \right)^t
\]
\[
\leq n^\alpha \left( \frac{\log n}{n^\alpha} \right)^{t/2},
\]
where for the last inequality we used our assumption that \( \frac{k^2}{K} \sim \frac{\log n}{n} \), which implies that \( \frac{k^2}{K} = \left( \frac{n^\alpha}{K} \right)^2 = \frac{\log n}{n^\alpha - 1} \). Thus, \( \Pr[\mathcal{E}(i)] \leq \frac{\log n}{n} = o(1) \).

**Theorem 4.5:** Assume \( K := n^\alpha \) with \( \alpha \geq 2 \). Then, with high probability, the kryptograph is unassailable.

**Proof:** By Lemma 4.3, with high probability, the graph has \( \Theta(n \log n) \) edges. Let \( an \log n \) be the number of edges, where \( a > 0 \) is a constant. By Lemma 4.4, with high probability, no bin contains more than \( 5 \) balls, i.e. every key is used at most by \( \frac{n}{5} \) edges. Thus, in order to capture \( a n \log n \) edges (i.e. a constant fraction), the adversary must capture at least \( a n \log n / 10 \) useful bins. By Lemma 4.2, with high probability, no vertex can contribute more than \( 10 \log n \) useful bins. It follows that, with high probability, the adversary must capture at least \( a n \log n / 100 \log n = \Theta(n) \) vertices.

**V. UNSPLITTABLE NETWORKS**

We now show that the adversary cannot easily partition the network into two linear size sets, corrupting all edges between them.
Given a set of vertices $S$ we denote by $(S, V-S)$ the set of edges with one endpoint in $S$ and the other in $V-S$, i.e. the set of edges that cross the cut. We will show that, with high probability, simultaneously for all $S$ with linearly many nodes, the number of edges that cross the cut is $\Theta(n \log n)$. The results of Section IV imply that in order to corrupt all these edges the adversary must capture a linear fraction of the nodes, i.e. the network is unsplittable.

Again, we will show the result in the full visibility case under the hypothesis $K = n^{-\alpha}$ with $\alpha \geq 2$. The result holds in the general case if $K = \Omega(n \log n)$ but the somewhat lengthy proof is omitted.

**Lemma 5.1:** With high probability, for all $S$ of size in the range $[0.1, 0.5n]$, $(S, V-S)$ contains $\Omega(n \log n)$ edges.

**Proof:** Fix a set $S$ of size $s \in [0.1n, 0.5n]$. To show that there are $\Omega(n \log n)$ edges leaving $S$, we proceed indirectly. Consider the set of triples

$$T = \{(v, w, g) : v \neq w \text{ and } g \in K_r \cap K_w\}.$$

Let $T(S) = \{(v, w, g) \in T : v \in V \text{ and } w \in V-S\}$. We show that (a) with high probability, for all $S$, $|T(S)| = \Omega(n \log n)$, and (b) with high probability for every pair $(v, w)$, there are at most three keys $g$, such that $(v, w, g) \in T$. Thus, the number of edges in $(S, V-S)$ is at least $|T(S)|/3 = \Omega(n \log n)$. As in the proof of Lemma 6.4, we will condition on the event $\mathcal{E}(S)$. Note that $|T(S)| \geq \sum_{v \in V-S} |k(S) \cap K_r|$, and

$$E\left[\sum_{v \in V-S} |k(S) \cap K_r| \right] = \frac{|k(S)| k(n-s)}{K} \geq \frac{sk^2(n-s)}{10K} \geq c_0 n \log n,$$

for some constant $c_0 > 0$ (independent of $|S|$). (For the last inequality we used $\frac{k^2}{K} \sim \frac{\log n}{n}$ and $0.1n \leq s \leq 0.9n$.) We can use Lemma 3.7 (with $p \leftarrow \frac{k}{K}$, $\ell \leftarrow n-s$ and $a \leftarrow |k(S)|$) and a routine computation, to conclude that the probability that $|T(S)|$ is less than $(c_0/2 \log n)$ is at most exp$(-(c_0/8)\log n)$. Thus,

$$\Pr(|T(S)| < (c_0/8)\log n) \leq \Pr[\mathcal{E}(S)] + \exp(-(c_0/8)\log n).$$

Using the union bound we obtain (by summing over all $S$ such that $\beta n \leq |S| \leq 0.5n$),

$$\Pr[\exists S : \beta n \leq |S| \leq 0.5n, |T(S)|] \leq \sum_S [\Pr[\mathcal{E}(S)] + \exp(-(c_0/8)\log n)] \leq 2^{-\Theta(n)},$$

where we use Lemma 6.2 to justify the last inequality. This completes part (a) of our argument.

For part (b), fix a pair of distinct vertices $(v, w)$. If $|K_r \cap K_w| \geq 4$, then there is a subset $J \subseteq K_r$ of size 4 such that $J \subseteq K_w$. Thus, by the union bound again

$$\Pr(|K_r \cap K_w| \geq 4) \leq \binom{k}{4} \binom{K-4}{k-4} (\binom{K}{k})^{-1} \leq O\left(\left(\frac{k^2}{K}\right)^4\right).$$

Since, $\frac{k^2}{K} \sim \frac{\log n}{n}$, this quantity is $O(n^{-3})$. There are at most $n^2$ pairs $(v, w)$, so with probability $1-o(1)$, for all $(v, w)$ there are at most three keys $g$ such that $(v, w, g) \in T$. This establishes part (b) and completes the proof of the lemma.

**Theorem 5.2:** Assume $K := n^\alpha$ with $\alpha \geq 2$. Then, with high probability, the network is unsplittable.

**Proof:** Follows immediately from Lemma 5.1 and Theorem 4.5.

VI. THE EMERGENCE OF A GIANT COMPONENT

In this section, we prove that if $\frac{k^2}{K} \sim \frac{C_0}{n}$, for an appropriately large constant $C_0$, then with high probability the kryptograph has a giant connected component. [To keep the proofs simple, we do not attempt to optimize the constant.] In particular, we will show that with high probability there exists a connected component of size $0.9n$. The same argument can be used to show the existence of a connected component of size $\gamma n$, for any $\gamma \in (0, 1)$.

Let $k(S)$ denote the set of keys chosen by vertices of $S$, and let $s$ denote the size of a set $S$. The expected size of $k(S)$ is roughly $ks$. The next lemma shows that it is extremely unlikely that any large (linear size) set has a combined key ring that is much smaller than this. From now on, let

$$\alpha, \beta := \frac{1}{10}.$$

**Definition 6.1:** For a set $S \subseteq V$, let $\mathcal{E}(S)$ denote the event $k(S) < \alpha k|S|$.

In the next lemma we assume $\frac{k^2}{K} = \frac{c}{n}$ where $c \leq k$. In this way the lemma holds for the case $\frac{k^2}{K} \sim \frac{1}{n}$, which is relevant for this section, as well as the case $\frac{k^2}{K} \sim \frac{\log n}{n}$, which is relevant for the next section.

**Lemma 6.2:** Let $\frac{k^2}{K} = \frac{c}{n}$ where $c \leq k$. Then,

$$\sum_{S \subseteq V : \beta n \leq |S| \leq \frac{2}{c}} \Pr[\mathcal{E}(S)] = 2^{-\Theta(n)}.$$

**Proof:** We first bound $\Pr[\mathcal{E}(S)]$ for a fixed set $S$, and then sum over all choices of $S$. We carry out our calculations pretending that the vertices obtain their key rings by sampling with replacement; note that this only makes or claim stronger. Fix a set $S$, and let $s = |S|$. By
Now, we sum over all choices $S$ (with $\frac{n}{10} \leq |S| \leq \frac{n}{2}$),
\[
\sum_S \Pr[\mathcal{E}(S)] \leq \left( \sum_{s=\lfloor \frac{n}{10} \rfloor}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{s} e^a \left( \frac{\alpha s}{n} \right)^{1-a} \right)^{sk} \leq \left( \sum_{s=\lfloor \frac{n}{10} \rfloor}^{\lfloor \frac{n}{2} \rfloor} \left( \frac{en}{s} \right)^s e^a \left( \frac{\alpha s}{n} \right)^{1-a} \right)^{sk} \leq \left( \sum_{s=\lfloor \frac{n}{10} \rfloor}^{\lfloor \frac{n}{2} \rfloor} \left( e^{\alpha k+1} a^{(1-a)k} \right) \left( \frac{s}{n} \right)^{1-a} \right)^{sk} \leq \frac{1}{2^k} = o(1).
\]

For the last inequality we used our assumption $k \geq 2$, which implies that
\[
e^{\alpha k+1} a^{(1-a)k} \leq e^k a^{(1-a)k} \leq (e^a)^{1-a} < 1.
\]

The next lemma says that with high probability all large enough subsets of nodes must be connected to a vertex outside of the set. This easily implies the existence of a giant component.

**Definition 6.3:** A cut is a non-empty subset $S \subseteq V$ of size at most $\frac{n}{7}$ such that there is no edge between $S$ and $V - S$. Let $\mathcal{E}(S)$ denote the event “$S$ is a cut”.

**Lemma 6.4:** Let $\frac{1}{2} \leq \frac{c}{k}$ where $c \leq k$. Then,
\[
\Pr[\exists S, \beta n \subseteq |S|, \mathcal{E}(S)] \leq 2^{-\beta n}.
\]

**Proof:** Fix a set $S$ of size $s$, such that $\beta n \leq s \leq \frac{n}{2}$.

Then,
\[
\Pr[\mathcal{E}(S)] \leq \Pr[\mathcal{E}(S)] + \Pr[\mathcal{E}(S) \mid \neg \mathcal{E}(S)].
\]

We have already analyzed the first term in the previous lemma. Let us focus on the second term:
\[
\Pr[\mathcal{E}(S) \mid \neg \mathcal{E}(S)] \leq \left( 1 - \frac{\alpha s}{K} \right)^{|V-S|k} \leq \left( 1 - \frac{\alpha s}{K} \right)^{kn/2} \leq \exp \left( -\frac{\alpha sk^2}{2K} \right) \sim \exp \left( -\frac{C_0 \alpha s}{2} \right) \ll 4^{-n}.
\]

[For the last inequality we choose $C_0$ large enough.]

Using the union bound (summing over choices of $S$ such that $\beta n \leq |S| \leq \frac{n}{2}$), we have
\[
\Pr[\exists S, \beta n \subseteq |S|, \mathcal{E}(S)] \leq \sum_{S \subseteq V} \Pr[\mathcal{E}(S)] \leq \sum_{S \subseteq V} \Pr[\mathcal{E}(S)] + \Pr[\mathcal{E}(S) \mid \neg \mathcal{E}(S)].
\]

The claim now follows from Lemma 6.2 and (7).

**Theorem 6.5:** With high probability, the kryptograph has a connected component of size at least $0.9n$.

**Proof:** Assume the event of the statement of Lemma 6.4. Thus, if $S$ is a cut then it has fewer than $0.1n$ vertices. In particular, the graph has no connected component with size in the range $[0.1n, 0.9n]$. Now, if the largest connected component has more than $0.9n$ vertices, then we are done. Otherwise, it must have fewer than $0.1n$ vertices. Start collecting components $S_1, S_2, \ldots$ until their total size exceeds $0.1n$. The total size is at most $0.2n$, and we get a cut with more than $0.1n$ vertices, a contradiction.

Just similarly to what has been done for the connected kryptograph, when Condition 2 holds the resulting giant connected component can be shown to be unassailable and unsplitable. The proof is slightly more technical and it is omitted for the sake of brevity.

**VII. Experiments**

We set up a number of experiments to back up our theoretical findings. A first experiment has been done to check that, if parameters are chosen according to Condition 1, then the sensor network is actually connected with high probability (note that this first part of the experiment replicates what has been done in [6]), and, more importantly, that if parameters are chosen according to Condition 2, then a giant connected component emerges in the network with high probability.

The details of the experiment are as follows: Network size $n$ ranges from 1,000 to 8,000 sensors, communication radius is set to $0.1$, pool size to $n \log n$, and key ring size to $c \log n$, where $c$ is set to 10. Constant $c$ depends on the communication range and must be experimentally tuned. With these parameters, the network is “almost always” completely connected, meaning that full connectivity fails with probability less than $0.01$ when $n$ is 1,000, and smaller and smaller probability for larger networks (when $n=2,000$ and larger, for example, we have got no disconnected network in the experiments after more than 10,000 runs). Note that probability of less than $0.01$ is ok for our purposes, it is enough to slightly increase constant $c$ to make it as small as desired.

To check the existence of a giant connected component with much smaller key rings (asymptotically smaller, as shown theoretically), we performed the same experiment setting the key ring size in accordance with Condition 2. To help visualize what a giant connected
component looks like in a kryptograph, in Figure 1 it is shown a sensor network where gray edges represent physical links and black edges secure links. Secure links in Figure 1 form a giant connected component. The same network is shown in Figure 2 where gray edges are secure links while black edges are those in the giant connected component.

Experimentally, a key ring of size \(11 \sqrt{\log n}\), same pool size of \(n \log n\), is enough to guarantee the emergence with high probability of a giant connected component of at least 90% of the sensors for \(n\) ranging from 1,000 to 8,000. Constant 11 depends on communication range and on the size of the desired giant component. Indeed, further experiments show that a key ring of size \(13 \sqrt{\log n}\) guarantees a giant component of 95%, while a key ring of size \(15 \sqrt{\log n}\) guarantees a giant component of 99% for the same network sizes. These conclusions have been validated by thousands of runs. See Figure 3 to appreciate the large gap between the key ring size needed for full connectivity and the key ring size for the presence of a giant component. This gap results in different structural properties of the kryptograph, for example giant components are much sparser. Figure 4 shows the average number of edges of the networks for key ring size of \(10 \log n\), \(11 \sqrt{\log n}\), \(13 \sqrt{\log n}\), and \(15 \sqrt{\log n}\); that is, in the case of connectivity, and in the presence of giant components of 90%, 95%, and 99% of the network sensors.

Let’s now turn to security. All the following experiments have been performed with the parameter choice tuned in the experiments above. It is important to realize there may be keys in the pool that are used (individually) to protect many links, more than a constant number. These keys are more “important”. This is a problem for security since the omniscient adversary knows these keys, and in which key rings they are stored, and can perform the attack in such a way to collect them. Since we know from our theoretical results that the adversary still cannot get a big advantage by using the optimal strategy, it must be true that important keys are few in the kryptograph. To experience this phenomenon, we have
performed an experiment to draw the outline of key load for all the keys in the pool. Figure 5 shows how many links are secured using the same individual key, for every key in the pool starting from the most important to the least important. The experiment has been done with the above described choice of parameters to get connectivity in a 8,000 node network. A large fraction of the keys are not used at all, while there are indeed keys that are used to secure as many as 14 links. However, it is also evident from the outline that there are not too many important keys in the network. Figure 6 zooms on the first 100 keys.

To see whether our first intuition is correct (the outline can vary considerably when \( n \) changes) with set up an experiment to check whether the kryptograph and the giant components are unsplittable. We assume that the adversary chooses to split the network into two parts of roughly equal size, by vertically cutting the geographical area into two parts. Figure 8 and Figure 9 show what this kind of cut looks like in a connected kryptograph and in a giant connected component. As shown in Figure 10, again, the number of important keys that must be collected to perform the cut grows linearly with network size, for the whole connected kryptograph and for every giant component. That is, connected kryptographs and giant connected components in kryptographs are both unassailable and unsplittable, even in networks of practical sizes.

**References**

Fig. 8. Vertical geographical cut to split a connected kryptograph into two parts of roughly the same size. Black edges are those that cross the cut.


