Connectivity Properties of Secure Wireless Sensor Networks

ROBERTO DI PIETRO, LUIGI V. MANCINI, ALESSANDRO MEI,
ALESSANDRO PANCONESI
University of Rome "La Sapienza", Italy
and
JAIKUMAR RADHAKRISHNAN
Tata Institute of Fundamental Research, Bombay, India

We address the problem of connectivity in Secure Wireless Sensor Networks (SWSN) using random pre-distribution of keys. First, we propose a geometric random model for SWSNs. Under this new and realistic model, we describe how to design secure and connected networks using a small constant number of keys per sensor. Moreover, we formally prove that these networks can tolerate the presence of $o(N/\log N)$ malicious and colluding nodes, where N is the size of the network. This is, to the best of our knowledge, the first asymptotic result on secure WSN resilience in the framework of random pre-distribution of keys. Lastly, we report extensive simulations to support the above stated results and demonstrate how connectivity can be guaranteed for a wide interval of practical network sizes and sensor communication ranges.

Categories and Subject Descriptors: []:

1. INTRODUCTION

A Wireless Sensor Network (WSN) is a collection of sensors whose size can range from a few hundred sensors to a few hundred thousand or possibly more. The sensors do not rely on any pre-deployed network architecture, thus they communicate via an ad-hoc wireless network. The power supply of each individual sensor is provided by a battery, whose consumption for both communication and computation activities must be optimized. Distributed in irregular patterns across remote and often hostile environments, sensors should autonomously aggregate into collaborative, peer-to-peer networks. Sensor networks must be robust and survivable in order to overcome individual sensor failure and intermittent connectivity (due, for instance, to a noisy channel or a shadow zone).

It is widely believed that WSNs can be useful in a plentiful variety of settings. In many applications establishing secure pair-wise communications can be useful. In particular, it is a pre-requisite for the implementation of secure routing, and can be useful for secure group communications as well. However, due to the scarceness of resources, public key cryptography is not a viable solution; confidentiality has to be enforced by using symmetric key algorithms. Key management is thus a central issue in secure wireless sensor networks.

This work was partially supported by the WEB-MINDS project from the Italian MIUR under the FIRB program, by the EU under the IST-2001-34734 EYES project, and by the Italian MIUR under the "Progetto Giovani Ricercatori".

In key pre-distribution schemes, the symmetric keys are distributed to all sensors before deployment on the ground. If we know which sensors will be in the same neighborhood before sensor deployment, keys could be assigned to sensors a priori using this information. However, such a knowledge does not often exist, as assumed by a number of the key pre-distribution schemes in the literature [Carman et al. 2000; Di Pietro et al. 2003; Basagni et al. 2001] and by the random key pre-deployment scheme [Eschenauer and Gligor 2002], one of the most promising solutions recently proposed.

A *Secure Wireless Sensor Network* (SWSN) is composed on N sensors. Each sensor is pre-assigned a *key ring* of k secret keys randomly drawn from a common pool of K random keys. The sensors are then randomly deployed in a given geographical area. Two sensors share a secure communication link if they lie within communication range and they share a common pre-assigned key. A fundamental problem in secure wireless sensor networks is to choose proper k and K such that the network is connected by using secure links alone. This problem has been earlier addressed by Eschenauer and Gligor in [Eschenauer and Gligor 2002]. Their methodology is based on the well-known random graph model and on the classical result on connectivity from Erdös and Rényi [Erdös and Rényi 1960]. However, we argue that the model in [Eschenauer and Gligor 2002] is not completely satisfactory for secure wireless sensor networks. The model does not include a notion of geometric position of the sensors in the space and, as a consequence, it cannot properly model the inherent locality of physical visibility among sensors. In this respect, our model is more realistic, since it does take into consideration the physical position of the sensors. More seriously perhaps, the Erdös—Renyi model assumes that edges exist independently. As shown in this paper this is far from being true and therefore the Erdös—Renyi model cannot be used as a reliable guide to design secure wireless networks.

In this paper, we show how to design secure wireless sensor networks that are connected with high probability even if each sensor is assigned a small constant number of keys. Our model gives precise indications on how to choose the relevant parameters, pool size and key ring size, in order to have very high probability of connectivity. In particular, we prove that key rings of very small size, two or more, already ensure high probability of connectivity. These model predictions are fully confirmed by our extensive set of experiments. Moreover, under connectivity properties, we formally prove that any attacker willing to compromise at least a constant fraction of the network by capturing sensors at random and collecting all the keys they store must compromise at least $\Omega(N/\log N)$ sensors of the network. This result is at most a factor of $\log N$ far from optimality. To the best of our knowledge, this is the first asymptotic bound on the resilience of wireless sensor networks using random pre-distribution of keys. Lastly, we support the above mentioned theoretical results with extensive simulations. The experiments compel the conclusion that our design guarantees connectivity and resilience for a wide interval of practical network sizes and communication ranges.

2. RELATED WORK

The idea of probabilistic key sharing for WSNs is introduced by Eschenauer and Gligor [Eschenauer and Gligor 2002]. The authors also provide a simple and centralized algorithm for re-keying in a distributed WSN. Later, in [Chan et al. 2003], three mechanisms are described in the framework of random key pre-distribution. First of all, the *q-composite* random key pre-distribution scheme, a modification of the basic scheme in [Eschenauer and Gligor 2002], achieves better security under small scale attack while trading off increased vulnerability in the face of a large scale physical attack on the network sensors. Secondly, the multi-path key reinforcement protocol substantially increases the security of the channel by leveraging the security of other links. Lastly, the random-pairwise keys scheme assigns private pairwise keys to randomly selected pairs of sensors so as to guarantee that the rest of the network remains fully secure even when some of the sensors have been compromised. Moreover, this latter scheme supports node to node authentication.

Two recent schemes build up a secure pairwise channel which combine a deterministic technique with a pre-distribution random scheme. The first scheme is proposed in [Du et al. 2003]. The authors use a deterministic protocol proposed by Blom [Blom 1985] 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 λ-*secure property*: as long as no more than λ nodes are compromised, the network is perfectly secure. A λ-secure data structure built this way is called a key space. The authors in [Du et al. 2003] create a set $\mathcal{W}$ composed of ω key spaces, and randomly assign up to τ 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 [Liu and Ning 2003]. In principle, this work is similar to [Du et al. 2003], where Blundo et al's polynomial scheme [Blundo et al. 1993] is used instead of Blom's.

Connectivity properties have been studied for non-secure wireless sensor networks as well. In [Bettstetter 2002], a geometric random model has been used to investigate minimum node degree and h-connectivity. Using a recent asymptotic result from Penrose [Penrose 1999], Bettstetter experimentally shows how to compute a communication range r such that, 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.

3. CONNECTIVITY OF SECURE WIRELESS SENSOR NETWORKS

3.1 Preliminaries

We say that $f(n) = o(1)$ if $f(n)$ goes to zero as n goes to infinity. If an event (depending on n) happens with probability $1 - o(1)$, we say that it occurs with *high probability*.

FACT 1. (UNION BOUND) *Let $E_1, \ldots, E_m$ be m events. Then,*

$$\Pr\left[\bigcup_{i=1}^{m} E_i\right] \leq \sum_{i=1}^{m} \Pr\left[E_i\right].$$

FACT 2. (CHERNOFF-HOEFFDING BOUNDS) *Let $X = \sum_{i=1}^{n} X_i$ where the X_i's are independently distributed in* $[0,1]$. *Then,*

$$\Pr[X < EX - t] \leq e^{-2t^2/n}.$$

We recall some basic facts and definitions from graph theory (see for instance [Bollobás 1979]). As customary $V(G)$ and $E(G)$ denote the vertex and the edge set of a graph G, respectively. Given a graph $G = (V, E)$ a *cut* is a proper subset $S \subseteq V$ such that there is no edge connecting a vertex in S with a vertex in $V - S$.

FACT 3. *A graph G is connected if and only if it has no cuts.*

Let S be a set. The collection of all subsets of S of cardinality k is denoted as

$$\binom{S}{k}.$$

The terms point, node and vertex will be used interchangeably.

3.2 Connectivity

We want to study the connectivity properties of the following geometric, random graph model.

Definition 1. Let K be the size of a finite set of keys, and let $k \le K$ 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 $G^N_{r,k,K}$ 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*, whose indexes are in $K_u \subseteq [K]$ by sampling $[K]$ with 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 $G^N_{r,k,K}$ is called a **kryptograph** with parameters r, k, K and N. In the special case in which every two nodes are within transmission range, the so-called *full visibility* case, the resulting graph is denoted as $G^N_{k,K}$.

In the sequel, for sake of simplicity we shall identify $[K]$ with the set of keys and K_u with the key ring of a vertex u. It is important to realize that edges of a kryptograph do not exist independently. Consider for instance three points x, y and z, all within distance r of each other, whose key rings are of size 2. And suppose that $K = 10^4$. This is a realistic scenario since, as we shall see, key rings of size 2 suffice for high probability of connectivity. Assume that we know already that edges xy and yz exist. What is the probability that edge xz also exists? If we assume independence then this probability is $\sim \frac{1}{5000}$, but in reality $\Pr[xz \text{ exists} \mid \text{both } xy \text{ and } yz \text{ exist}] \sim \frac{1}{2}$!

This clearly shows that the classical Erdös-Renyi model proposed in [Eschenauer and Gligor 2002] is not a reliable guide to design secure wireless networks, the geographical nature of the network notwithstanding.

Assuming that they key rings are generated by sampling with replacement not only simplifies the analysis of connectivity. In fact, sampling without replacement can only be better. To see this, suppose each node picks a set of size k in the following way. It first picks a set by sampling with replacement k times. Now, if did not pick k distinct elements it picks whatever more is needed by sampling without replacement. So, in the end it has a set of size exactly k, and the distribution of this key ring is uniform. Thus, we can always assume that key rings sampled without replacement were generated by this process, but the key rings we consider in the proofs (i.e. the first k samples) are actually subsets of the actual sets the nodes hold. So, if there is connectivity using sampling with replacement there must be connectivity using sampling without replacement.

Generating key rings without replacement on the other hand slightly worsen the security of the network, since the key rings on average are slightly smaller. As we shall see however the effect is a completely negligible. Thus, it appears that the kryptograph is the right graph-theoretic model for SWSNs.

As we shall see below, in the full visibility case, where every two sensors are within transmission range, in order to have high probability of connectivity in kryptographs (and thus in secure wireless sensor networks), it suffices to have $k \geq 2$ and $K \approx N/\log N$. In the general case, when we distribute N sensors in the unit square and r is the transmission radius, the latter becomes $K \approx r^2 N/\log N + 2\log r$. To increase the probability of connectivity, we can either increase k or decrease K. The condition on K is quite robust, i.e. values of K that are bigger of a constant factor than the above bounds can in practice guarantee high probability of connectivity, as shown by the experiments in Section 5. This has important (positive) consequences for the security of the network. Similarly, the very weak condition on k ($k \geq 2$) allows for great flexibility.

We introduce some notation and definitions. We divide the unit square into ℓ^2 square cells of equal size, where ℓ is the smallest integer greater than or equal to $\sqrt{5}/r$. This choice ensures that any two points in adjacent cells, or in the same cell, are within transmission radius.

Given a collection $\mathcal{C} = \{S_1, \ldots, S_n\}$ of key rings, the graph $H := H(\mathcal{C})$ is defined as follows. The vertex set of H is $[K]$, the set of all keys. A pair of keys xy is an edge of H if there exists a set $S \in \mathcal{C}$ to which both x and y belong. The crux of our argument is the following. A collection of key rings $\mathcal{C}$ induces two different graph. The kryptograph and the the graph $H(\mathcal{C})$ just defined. If the collection of key rings is chosen according to definition 1, then H is connected with very high probability and if H is connected, so is the kryptograph. Let us now proceed formally.

Definition 2. A collection $\mathcal{C} = \{S_1, \ldots, S_n\}$ of key rings is a *good collection* if $H(\mathcal{C})$ is connected.

A set P of points such that every two of them are within transmission range is a set of *close neighbours*.

Definition 3. Let P be a set of n close neighbours, and let $\mathcal{C}$ be the corresponding collection of key rings. The graph $G_{P,\mathcal{C}}$ is the kryptograph whose vertex set is P and where uv is an edge if $K_u \cap K_v \neq \emptyset$, where $K_u, K_v \in \mathcal{C}$.

LEMMA 1. *Let P be a set of close neighbours and let $\mathcal{C}$ be the corresponding set of key rings. If $H(\mathcal{C})$ is connected then $G_{P,\mathcal{C}}$ is connected.*

PROOF. By Fact 3 if we show that $G := G_{P,\mathcal{C}}$ has no cuts the claim follows. Let S be a proper subset of $V(G)$, let $k(S) := \cup_{u \in S} K_u$ and let $y \in V - S$. Now, since $H := H(\mathcal{C})$ is connected, the set $k(S)$ is not a cut for H. Therefore there must be an edge between a vertex (key) $x \in k(S)$ and a vertex (key) $y \in [K] - k(S)$. By definition of H, xy is an edge of H only if there is a key ring K_v such that $x \in K_v$ and $y \in K_v$. By definition of $k(S)$, $v \notin S$. Now, since $x \in k(S)$ there must be some other key ring K_u such that $x \in K_u$ and $u \in S$. But this implies that uv is an edge of G and therefore S is not a cut of G. $\square$

The following theorem shows that key rings of size $k \geq 4$ suffice for very high probability of connectivity, provided that $K = n/\log n$. The case $k \geq 2$ also holds, but it must be dealt with separately.

THEOREM 1. *Let P be a set of n close neighbours and let $\mathcal{C}$ denote the corresponding set of key rings generated by sampling k times without replacement from a set $[K]$. Let $c \geq 1$, $k \geq 2(c+1)$ and $K := n/\log n$. Then, the probability that $\mathcal{C}$ is not a good collection is at most*

$$b(n) := \frac{n^{-c}}{1-n^{-c}} \sim n^{-c}.$$

PROOF. We need to show that the probability that $H := H(\mathcal{C})$ is not connected is at most $b(n) \sim n^{-c}$. We will do this using Fact 3. Fix a subset S of $[K]$ of size $s \leq K/2$. We say that S is *crossed* if there exists a key ring $K_i \in \mathcal{C}$ that non-trivially intersects both S and the complement of S. The probability that S is not crossed is,

$$\begin{aligned}
\Pr[S \text{ is not crossed}] &\leq \left(\left(\frac{s}{K}\right)^k + \left(1-\frac{s}{K}\right)^k\right)^n \\
&\leq \left(1-\frac{s}{K}\right)^{kn}\left(1+\left(\frac{s}{K-s}\right)^k\right)^n \\
&\leq \exp\left\{-\frac{skn}{K}\right\}\exp\left\{\left(\frac{s}{K-s}\right)^k n\right\} \\
&\leq \exp\left\{-\frac{skn}{2K}\right\}\exp\left\{\left[\left(\frac{s}{K-s}\right)^k - \frac{sk}{2K}\right]n\right\}.
\end{aligned}$$

Since $s \leq K/2$, we have that $s/(K-s) \leq 1$. So, we can ignore the exponent k above and bound the failure probability by,

$$\Pr[S \text{ is not crossed}] \leq \exp\left\{-\frac{skn}{2K}\right\}\exp\left\{\left[\frac{s}{K-s} - \frac{sk}{2K}\right]n\right\}.$$

Again, $K - s \geq K/2$, so that

$$\Pr[S \text{ is not crossed}] \leq \exp\left\{-\frac{skn}{2K}\right\}\exp\left\{\left[\frac{2s}{K} - \frac{sk}{2K}\right]n\right\} \leq \exp\left\{-\frac{skn}{2K}\right\}.$$

The last inequality follows from the assumption $k \geq 4$ which implies that the second exponential is no more than 1. Recalling that $k \geq 2(c+1)$ and that $K := n/\log n$, we can bound the probability that H is disconnected using the union bound, as follows.

$$
\begin{aligned}
\Pr[H \text{ disconnected} &= \Pr[\exists S, S \text{ is a cut}] \\
&= \Pr[\exists S, S \text{ is not crossed}] \\
&\le \sum_{s=1}^{K/2} \sum_{S \in \binom{[K]}{s}} \Pr[S \text{ is not crossed}] \\
&< \sum_{s=1}^{\infty} n^s \Pr[S \text{ is not crossed}, |S| = s] \\
&\le \sum_{s=1}^{\infty} n^s \exp\left\{-\frac{skn}{2K}\right\} \\
&\le \sum_{s=1}^{\infty} n^{-cs} \\
&= \frac{n^{-c}}{1-n^{-c}} \sim n^{-c}
\end{aligned}
$$

This concludes the proof of Theorem 1. □

Corollary 1. *Let P be a set of n close neighbours and let $\mathcal{C}$ denote the corresponding set of key rings generated by sampling k times without replacement from a set K. Let $k \ge 2$ and $K := n/\log n$. Then, the probability that $\mathcal{C}$ is not a good collection is a*

$$o(1) \sim \frac{1}{\log n}.$$

Proof. The proof is identical to that of the theorem, but the calculations differ. Let S be a subset of $[K]$ of size $s \le K/2$. Since $k \ge 2$, $K = n/\log n$ and $2(K-s) \ge K$, the probability that S is not crossed is,

$$
\begin{aligned}
\Pr[S \text{ is not crossed}] &\le \left(\left(\frac{s}{K}\right)^k + \left(1-\frac{s}{K}\right)^k\right)^n \\
&\le \left(\left(\frac{s}{K}\right)^2 + \left(1-\frac{s}{K}\right)^2\right)^n \\
&\le \left(1-\frac{2s(K-s)}{K^2}\right)^n \\
&\le \left(1-\frac{s}{K}\right)^n \\
&\le e^{sn/K} = n^{-s}.
\end{aligned}
$$

Thus,

$$\begin{aligned}
\Pr[H \text{ disconnected} &= \Pr[\exists S, S \text{ is not crossed}] \\
&\leq \sum_{s=1}^{K/2} \sum_{S \in \binom{[K]}{s}} \Pr[S \text{ is not crossed}] \\
&\leq \sum_{s=1}^{\infty} K^s n^{-s} \\
&= \sum_{s=1}^{\infty} \frac{1}{\log^s n} \\
&\sim \frac{1}{\log n}.
\end{aligned}$$

□

Theorem 1 and Corollary 1 show that in the full visibility case kryptographs are connected with high probability. If $k := 2(c+1)$ the probability that the graph is not connected is $\sim n^{-c}$. Corollary 1 gives a weaker, but still high probability guarantee for the case $k \geq 2$. We now extend the result to the general case.

First we establish another useful property of a good collection.

LEMMA 2. *Let P be the set of points inside a cell C, and let $\mathcal{C}$ be the corresponding collection of key rings. If $\mathcal{C}$ is a good collection, then any node in an adjacent cell is connected to some node inside C.*

PROOF. The claim follows by observing that if $\mathcal{C}$ is a good collection then $\cup_{S \in \mathcal{C}} S = [K]$. To see this, just observe that since $H(\mathcal{C})$ is connected, for all $x \in [K]$, $\{x\}$ is not a cut of H. □

Recall that by definition of ℓ a set of points inside a cell is a set of close neighbours. Theorem 1 says that inside each cell the graph is connected, while this lemma says that subgraphs contained in adjacent cells are connected to each other. This way we see that the whole of $G^N_{r,k,K}$ is connected. The next theorem simply estimates the probability that $G^N_{r,k,K}$ fails to be connected.

THEOREM 2. *Let $k \geq 2$ and $K := N/\log N$. Then, the probability that $G^N_{r,k,K}$ is not connected is $o(1)$.*

PROOF. Fix a cell C and let X be the random variable denoting the number of points to fall inside C. Then $EX = N/\ell^2$ and, by the Chernoff-Hoeffding bounds,

$$\Pr[\textit{there is a cell with less than } N/2\ell^2 \textit{ points}] \leq \ell^2 e^{-N/2\ell^4} = o(1).$$

Invoking Corollary 1 and using the union bound, we have a total error probability of at most,

$$\ell^2 e^{-N/2\ell^4} + \ell^2(1+o(1))\frac{1}{\log n} = o(1).$$

Therefore, $G^N_{r,k,K}$ is connected with probability $1 - o(1)$. □

4. RESILIENCE OF SECURE WIRELESS SENSOR NETWORKS

Kryptographs are introduced to model *secure* wireless sensor networks. Therefore, while connectivity is a fundamental property of SWSNs (we cannot really call it a network if it is disconnected), also fundamental is it to understand how resilient is a SWSN against external attacks to the communication confidentiality or against a collusion of a subset of the sensor nodes in the network.

We want to model the following attack: an external entity tampers with the sensors and collects all the keys in the key rings. The sensors to tamper with are chosen randomly among the ones yet to be compromised. The attacker's goal is to collect enough keys to decrypt as many network communications as possible. An equivalent attack can be carried on by a subset of the sensors in the SWSN that are actually malicious and cooperate to subvert the communication confidentiality by using all their knowledge and keys.

Definition 4. A *collusion* of nodes of kryptograph $G^N_{r,k,K}$ is a subset of the node set. Given a collusion of nodes, a key of the pool is compromised if it belongs to K_w, for some node w in the collusion. We also say that edge uv of kryptograph $G^N_{r,k,K}$ is compromised if and only if all the keys in $K_u \cap K_v$ are compromised.

In the following, we will assume that kryptographs are designed according to the guidelines in the previous section. Thus $K = \Theta(N/\log N)$ and the kryptographs are connected with high probability.

THEOREM 3. *Assume that a collusion of c nodes is uniformly chosen at random in a kryptograph. If the collusion can compromise a constant fraction of the network links with probability $\epsilon > 0$, then $c = \Omega(N/\log N)$.*

PROOF. Let a collusion of c nodes be uniformly chosen at random in $G^N_{r,k,K}$. Assume that $c = o(N/\log N)$ and that the collusion can compromise a constant fraction of the network links with probability $\epsilon > 0$. Since every node is assigned a constant number k of keys, the collusion as a whole has a collection of $o(N/\log N) = o(K)$ compromised keys. Consider an edge uv in $G^N_{r,k,K}$. Clearly, $K_u \cap K_v$ contains at least one key. Therefore, edge uv is compromised with probability $o(1)$. As a consequence, at most a fraction of $o(1)$ edges are compromised, on average. This directly implies that the collusion cannot compromise a constant fraction of the network links with probability ϵ. Since assuming $c = o(N/\log N)$ leads to a contradiction, it must be $c = \Omega(N/\log N)$. □

This result shows that secure wireless sensor networks can be designed such that they are connected and they are asymptotically "almost" perfectly secure against attacks and collusions, since the number of sensors to be compromised is almost of the order of the size of the whole network. In the following section, we will describe experiments that support this conclusion even for small network sizes.

5. SIMULATION RESULTS

In principle, it could be possible that our theoretical results do not exactly predict the properties of small sized secure wireless sensor networks. This is not the case. Our simulations say that, even though small networks are harder to connect (they might easily be physically disconnected), our results are robust enough to be valid

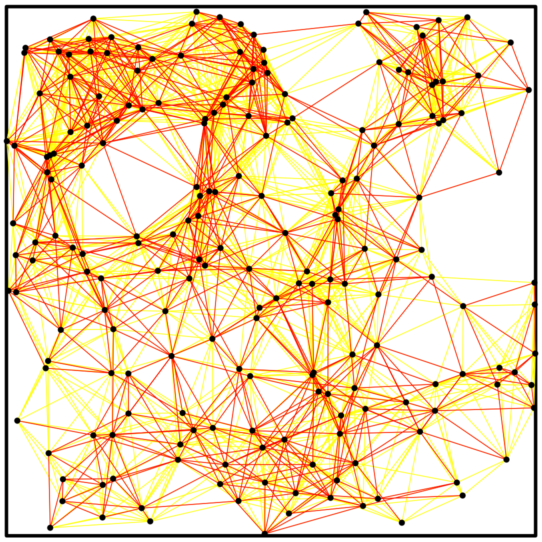

Fig. 1. Randomly generated secure sensor network of size 200. Communication range is 0.2, key ring size is 4, and pool size is 50. Lighter lines mean physical visibility, darker lines secure visibility. This graph is connected by using secure links alone.

starting from small network sizes. In all of the following experiments, key ring size is constant and fixed to 4. Clearly, if key ring size increases, connectivity probability also increases considerably.

To help understand the structure of secure wireless sensor networks, Figures 1, 2, and 3 show three similar networks where the pool size is increased from 50 to 100 and then to 150. Note that, as soon as the pool size is too big to guarantee connectivity, isolated sensors start to appear in the graph. This is perfectly analogous to what is predicted by well-known graph-theoretic results on similar random models and very important from a practical point of view. Indeed, even in the remote probability that our design methodology generate a disconnected graph, it is almost surely connected except a very small number of isolated points.

Figure 4 shows the probability of network connectivity as a function of the pool size. Network size N is 500 and 1,000, key ring size is 4, and communication range is 0.2. As it can be readily checked, a pool size $K = N/2 \log N$ guarantees that the sensor network be connected. Factor $1/2$ depends on constant k (set to 4 in our experiments) and on the communication range and is larger than the constant factor in our asymptotic results. This fact experimentally demonstrate that our theoretical results are robust. As a further experiment, with the same pool size $K = N/2 \log N$, we generated a large number of networks for increasing N and fixed communication range. These networks are virtually "always" connected. Indeed, we got no disconnected SWSNs among the 10,000 networks generated per each single size N from 500 to 10,000, step 100.

Lastly, we performed extensive experiments to validate Theorem 3. Figure 5 shows the number of sensors that a coalition must have to compromise 50% and 25% of the network, respectively. Both graphs can be almost exactly interpolated

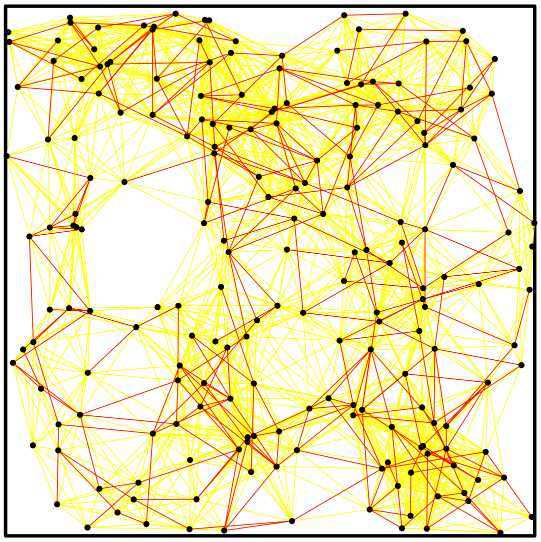

Fig. 2. Randomly generated secure sensor network of size 200. Communication range is 0.2, key ring size is 4, and pool size is 100. Lighter lines mean physical visibility, darker lines secure visibility. The network has a few isolated sensors.

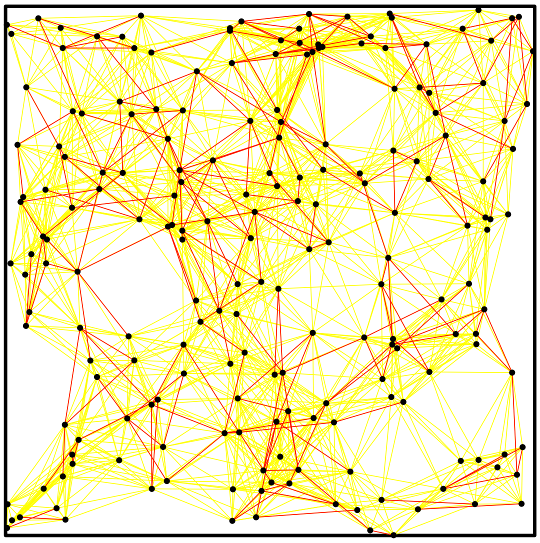

Fig. 3. Randomly generated secure sensor network of size 200. Communication range is 0.2, key ring size is 4, and pool size is 150. Lighter lines mean physical visibility, darker lines secure visibility. The network has a slightly larger number of isolated sensors and even some very small disconnected components.

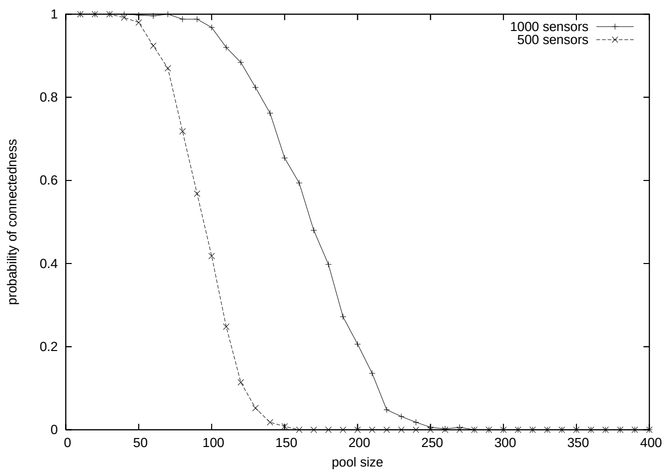

Fig. 4. Probability of connectivity of a secure sensor network of 500 and 1000 nodes as a function of the pool size.

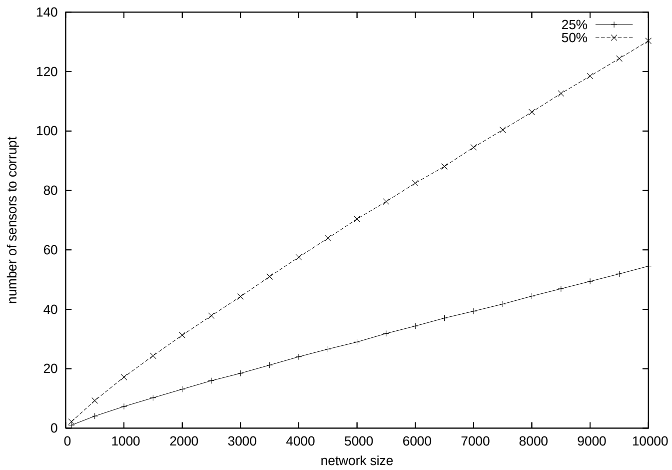

Fig. 5. Number of sensors that the attacker has to collect to compromise 50% and 25% of the network links. Pool size K is set to $N \log N$.

by function $cN/\log N$, where constant c depends on the fraction of network links to be compromised. This supports the practical soundness of our result on network resilience for all network sizes.

6. CONCLUDING REMARKS

In this paper we introduce a novel and realistic model for secure wireless sensor networks using random pre-distribution of keys. Under this model, the notion of local visibility is natural and connectivity can be formally studied. In particular, we show that pool size can be fixed in such a way that generated sensor networks are connected with high probability even assigning a small constant number of keys to each sensor. For example, our experiments shows that a SWSN of 10,000 sensors, communication range 0.2, key ring size 4, and pool size 550 is virtually always connected, in agreement with our theoretical results. Moreover, the SWSNs designed according to our guidelines prove to be almost perfectly secure against large scale attacks. In fact, to compromise a constant fraction of the network the attacker has to collect a number of sensors which is almost of the same order of the size of the whole SWSN. This is, to the best of our knowledge, the first asymptotic result on network resilience for secure wireless sensor networks.

REFERENCES

Basagni, S., Herrin, K., Bruschi, D., and Rosti, E. 2001. Secure pebblenets. In *Proceedings of the 2001 ACM International Symposium on Mobile ad hoc networking & computing, Long Beach, CA, USA. ACM Press.*

Bettstetter, C. 2002. On the minimum node degree and connectivity of a wireless multihop network. In *Proceedings of the* 3^{rd} *ACM international symposium on Mobile ad hoc networking and computing*. 80–91.

Blom, R. 1985. An optimal class of symmetric key generation systems. In *Advances in Cryptology: Proceedings of EUROCRYPT '84, volume 338 of LNCS.*

Blundo, C., Santis, A. D., Herzberg, A., Kutten, S., Vaccaro, U., and Yung, M. 1993. Perfectly-secure key distribution for dynamic conferences. In *Advances in Cryptology: Proceedings of CRYPTO '92, volume 740 of LNCS.*

Bollobás, B. 1979. *Graph Theory, an introductory course.* Springer-Verlag.

Carman, D. W., Kruus, P. S., and Matt, B. J. 2000. Constraints and approaches for distributed sensor network security. Tech. Rep. Technical Report 00-010, NAI Labs.

Chan, H., Perrig, A., and Song, D. 2003. Random key predistribution schemes for sensor networks. In *Proceedings of the IEEE Symposium on Security and Privacy.*

Di Pietro, R., Mancini, L. V., and Jajodia, S. 2003. Providing secrecy in key management protocols for large wireless sensors networks. *Journal of AdHoc Networks 1,* 4.

Du, W., Deng, J., Han, Y. S., and Varshney, P. K. 2003. A pairwise key predistribution scheme for wireless sensor networks. In *Proceedings of the* 10^{th} *ACM Conference on Computer and Communications Security (CCS '03).*

Erdös, P. and Rényi, A. 1960. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci. 5*, 17–61.

Eschenauer, L. and Gligor, V. D. 2002. A key-management scheme for distributed sensor networks. In *Proceedings of the* 9^{th} *ACM Conference on Computer and Communications Security (CCS '02).*

Liu, D. and Ning, P. 2003. Establishing pairwise keys in distributed sensor networks. In *Proceedings of the* 10^{th} *ACM Conference on Computer and Communications Security (CCS '03).*

Penrose, M. D. 1999. On k-connectivity for a geometric random graph. *Random Structures and Algorithms 15,* 2, 145–164.