Connected Component in Secure Sensor Network Induced by a Random Key Pre-Distribution Scheme

Bhupendra Gupta, and Subir Singh Lamba

***Abstract*—Wireless sensor network (WSN) has a wide range of applications in various areas. Many time the environment in which these sensor were deployed are hostile in nature and sensors have continuous attacks from the adversary, in such environmental conditions we need a secure communication between the sensors. For secure communication, neighbors must posses a secret common key or there must exists a key-path among these nodes. In this paper, the object of study is a random graph induced by the random key pre-distribution scheme of Eschenauer and Gligor under the assumption of full visibility. Here we establish the threshold value of the parameters (Key pool size and key-ring of an individual node) for which the entire network is almost surely a single connected component. We prove that for a network having N nodes, is a single connected component almost surely, if size of the key-ring is $m = \sqrt{2}\log\ N$ and the size of key pool is $K = N \log\ N$.**

***Index Terms*- Secure Sensor Networks, Random Key Pre-distribution Scheme, Secure Connectivity.**

I. Introduction

Wireless Sensor Networks (WSNs) is a distributed collection of sensor nodes. These wireless sensor networks have wide applications in various areas like in disaster recovery, military operation, tracking etc. Generally these sensor nodes having limited computational and communication capacities. Also WSNs do not have any pre-deployed network architecture (for example sensors like smart dust usually deployed by aeroplanes), some times it is due to the hostile environment. Thus sensors need to communicate via an ad-hoc wireless network. After deployment an individual sensor need to connect with the other sensors and create a secure connected wireless sensor network. Many applications (like in military operation) needs highly secure operation of sensor network, and have serious consequences if the network is compromised or disrupted. For this a secure pairwise communication must be established. Also the sensor network must be robust against the compromise of individual sensor due to failure and manipulation by an adversary.

One of the most promising approach for achieving a secure sensor wireless network is the random predistribution of keys introduced in [6].

We are initiated by describing the model introduced in [6].
A secured wireless sensor network is composed by a number (says N) of sensors. Each sensor is preloaded with a key-ring having m distinct encryption keys, randomly chosen from a pool of K keys. The shared key discovery take place during WSN initialization in the operational environment where every sensor discovers its neighbors and two sensors can form a secure link if they are within the wireless communication range and they have one or more common encryption key in their respective key rings.

In [6], authors raise a question that for a secure connected network, what should be the size of key pool and the size of the key-ring of an individual node? When every node have full visibility and mutually independent source link allocation. Here full visibility means that two node can communicate with each other, irrespective of their geographical position in the operational area. It might be a case when the communication range of sense deployed in the area is sufficiently lager than the operational area where they deployed.

Few extensions of the above described model has found in [3]. In [3], authors present three new schemes for key establishment using the framework of pre-distribution of a random key set among the node of network. The Schemes suggested in q-composite keys scheme, multipath-reinforcement scheme and

Manuscript received 21 April 2011.

B.Gupta is with the Indian Institute of Information Technology, Design & Manufacturing Jabalpur, MP, India 482005.E-mail: gupta.bhupendra@gmail.com, http://www.iiitdmj.ac.in/ bhupen

S.S.Lamba with Indian Institute of Information Technology, Design & Manufacturing Jabalpur, MP, India 482005.Email: subirs@gmail.com, http://www.iiitdmj.ac.in/ subirs

random pairwise keys scheme. After a detail analysis authors conclude that q-composite keys scheme achieves significantly improved security under small scale attack, while the (2-hop) multipath reinforcement scheme improves security at the cost of network communication overhead. Finally Authers [3], claim that the random pairwise scheme has the best security properties of the above three schemes.

In [7], author did the asymptotic analysis for key pre-distribution scheme for distributed sensor networks, and established relationship between interesting parameters like storage, connectivity etc. A new key pre-distribution scheme suggested in [2], this scheme allows different generation of sensors, i.e., sensors deployed at fixed time difference, to establish a secure communication. Here the sensor nodes have limited life time and refreshed periodically. using this scheme the network has self-healing ability, i.e., network recover its initial state in the next generation after the attack stop. Author shows that a RoK based network that is in the hostile environment is much less effective than a network having exiting key pre-distribution scheme. [8], presented two schemes to take advantage of sensors location information, aiming at improving pairwise key establishment in sensor networks. In first scheme, the closest pairwise keys scheme, is resistant to node capture attacks and has no limit on the total number of sensors. Its extended version further reduces the storage overhead and simplifies the dynamic deployment of new sensors. The second scheme is location-based key pre-distribution using bivariate polynomials, employs a threshold technique and provides a trade-off between the security against node capture.

A new identity based approach for key establishment between two neighbor nodes in wireless sensor networks initiated in [1]. This scheme supports dynamic node addition after initial deployment and also works for any deployment configuration. Another id-based random key pre-distribution scheme is introduced in [4], this scheme not only retains all the highly desirable properties of the schemes including high probability of establishing pairwise keys, tolerance of node compromise but also significantly reduces communication and computational costs of each node.

In [5], proves a threshold property which says when the number of compromised nodes is less than the threshold, the probability that any nodes other than these compromised nodes are affected is close to zero. Another interesting result for same network with full visibility has been derived in [10], authors shows the existence of zero-one law for the absence of isolated node in the network.

The present work is slightly inclined toward the analytic side, our main result is actually a strong law result. We derive expression of the size of key-ring of an individual node and size of key pool, for which the network is almost surly connected, i.e., the probability of connectivity is converges to 1 pointwise. Our result is parallel to the result in [9], where the authors suggest the pool size and the key-ring size for getting the connected network with high probability. The result in [9], is actually a weak law result suggested the values of parameters such that probability of existing a cut is approaches to 0, for sufficiently large N. While our result is a strong law result, and ensure the connectivity of the network in almost sure sense. Parameters suggested in our paper also support the results available in [9].

In this paper author proposed the proportion of key pool and key ring size for which the network can achieve connectivity with secure links. In section II, we define the basic terminology and supporting results. In section III we derive our main result which is a strong law result for the connectivity a sensor network with secure links and having random key pre-distribution scheme. expression for the probability of connectivity by 1-hop neighbor is established in section IV. Also we show that connectivity by 1-hop neighbor converges to 0 in probability.

II. Preliminaries.

We consider a 2-dimensional space, and a sequence $\mathcal{X}_N = \{X_1, X_2, \ldots X_N\}$, with N sensor nodes distributed uniformly over a compact space $C \subset R^2$.
Consider a key pool $\mathcal{K}$ having K encryption keys. Every node $X \in \mathcal{X}_N$ choose m different encryption key from $\mathcal{K}$, the key ring of X is denoted by $k(X) \subseteq \mathcal{K}$. Let $\mathcal{S} \subset \mathcal{X}_N$, then

$$k(S) := \cup_{X \in \mathcal{S}} k(X).$$

Our model is a random key graph G_N as suggested in [10], which as follows:

Definition 1: Let G_N be a graph with the vertex set $\mathcal{X}_N$ having N nodes and edge between the two node X_i and X_j represent at least one common encryption key in their respective key rings (since we assume the full visibility), that is edge set E is define as follows:

$$E := \{e_{ij} : k(X_i) \cap k(X_j) \neq \phi, \forall X_i, X_j \in \mathcal{X}_N\},$$

where e_{ij} is an edge between two arbitrary vertices X_i and X_j.

Also we assume full visibility for each node, i.e., each node can communicate with any other node if they have at least one common key in their key-rings.

Definition 2: Given a graph $G = (V, E)$, a cut is proper subset $S \subset V$ such that there is no edge connecting a vertex in S with any vertex of $V \setminus S$, in terms of graph theory cut S is a component of G.
Note that the $\mid S \mid \leq O(N)$, where $\mid S \mid$ is the size of cut $S \subset V$.

Definition 3: Let $f(x)$ and $g(x)$ be two functions defined on some subset of the real numbers. One writes

$$f(x) = O(g(x)), \qquad \text{as } x \to \infty,$$

if and only if, for sufficiently large values of x, $f(x)$ is at most a constant multiplied by $g(x)$ in absolute value. That is, $f(x) = O(g(x))$ if and only if there exists a positive real number M and a real number x_0 such that

$$\mid f(x) \mid \leq M \mid g(x) \mid, \qquad \forall\ x > x_0.$$

i.e.,

$$\left|\frac{f(x)}{g(x)}\right| \leq M, \qquad \forall\ x > x_0.$$

Definition 4: Let q_s be the probability that two arbitrarily chosen vertices does not have common keys in their respective key rings, and given by

$$\begin{aligned} q_s &= \frac{\binom{K}{m}\binom{K-m}{m}}{\binom{K}{m}\binom{K}{m}} \\ &= \frac{\binom{K-m}{m}}{\binom{K}{m}}. \end{aligned}$$

Then the probability that any two nodes share at least one key in their key-ring is given by

$$\begin{aligned} p_s &= 1 - q_s \\ &= 1 - \frac{\binom{K-m}{m}}{\binom{K}{m}}. \end{aligned}$$

Following lemma gives the lower and upper bounds of q_s, (the probability that two arbitrarily chosen vertices does not have common keys in their respective key rings).

Lemma 1: Let q_s is defined as above, then

$$\left(1 - \frac{m}{K-m}\right)^m \leq q_s \leq \left(1 - \frac{m}{K}\right)^m \leq \exp\left(-\frac{m^2}{K}\right);$$

and

$$\frac{m^2}{K} \leq 1 - q_s \leq \frac{m^2}{K-m}. \tag{1}$$

The proof of the lemma can be found in appendix.

III. Connectivity of Secure Sensor Networks.

The following theorem gives the critical values of size of key pool $\mathcal{K}$ and size of key-ring of an individual node, such that the network is connected.

Theorem 2: Let $S \subset V$ be a cut and let size of key pool, $K = \mid \mathcal{K} \mid = N \log N$ and size of key-ring an individual node X is $m = \mid k(X) \mid = c \log N$, where c is some positive constant. Then for sufficiently large N, and $c > \sqrt{2}$ network have a single component.

Proof. For this we consider

$$\begin{aligned} & P[\exists S, \mid S \mid = s, k(S) \cap k(V \setminus S) = \phi] \\ = \ & P[\exists S, \mid S \mid = s, \cup_{x \in S} k(x) \cap k(V \setminus \{x\}) = \phi] \\ \leq \ & \sum_{x \in S} P[\exists S, \mid S \mid = s, k(x) \cap k(V \setminus \{x\}) = \phi], \end{aligned} \tag{2}$$

by the Bole's inequality.

Now, consider probability that a particular vertex $x \in V$ is cut vertex. For being a cut vertex, the key ring of vertex x must not have any common key with the key ring of $V \setminus \{x\}$.

$$\begin{aligned} & P[k(x) \cap k(V \setminus \{x\}) = \phi] \\ = \ & \frac{\binom{K}{m}\left(\binom{K-m}{m}\right)^{N-1}}{\left(\binom{K}{m}\right)^N} = \left(\frac{\binom{K-m}{m}}{\binom{K}{m}}\right)^{N-1} \\ = \ & \left(\frac{(K-m)(K-m-1)\ldots(K-2m+1)}{K(K-1)\ldots(K-m+1)}\right)^{N-1} \\ = \ & \left(\left(1 - \frac{m}{K}\right)\left(1 - \frac{m}{K-1}\right)\ldots\left(1 - \frac{m}{K-m+1}\right)\right)^{N-1} \\ < \ & \left(1 - \frac{m}{K}\right)^{m(N-1)} < (\exp(-m))^{\frac{m(N-1)}{K}} \\ = \ & e^{-\frac{m^2(N-1)}{K}}. \end{aligned} \tag{3}$$

Substituting $m = c \log N$ and $K = N \log N$, in (3). Then we have

$$\begin{aligned} P[k(x) \cap k(V \setminus \{x\}) = \phi] &< e^{-c^2 \log N} \\ &= \frac{1}{N^{c^2}}. \end{aligned} \tag{4}$$

Using (4) in (2), we get

$$\begin{aligned} & P[\exists S, \mid S \mid = s, k(S) \cap k(V \setminus S) = \phi] \\ < & \sum_{x \in S} \frac{1}{N^{c^2}} = \frac{C}{N^{c^2-1}}, \end{aligned}$$

since the $\mid S \mid$ is of order of $O(N)$. For $c > \sqrt{2}$, the above probability is summable i.e.,

$$\sum_{N=1}^{\infty} P[\exists S, \mid S \mid = s, k(S) \cap k(V \setminus S) = \phi] < \infty.$$

Then by the Borel-Cantelli's Lemma , we have

$$P[\exists S, \mid S \mid = s, k(S) \cap k(V \setminus S) = \phi \quad i.o.] = 0.$$

Hence

$$k(S) \cap k(V \setminus S) \neq \phi \qquad \text{almost surly.}$$

This implies that probability that the key-ring $k(S)$ of an arbitrary subgraph $S \subset V$ have at least one common key with the $k(V \setminus S)$ the key-ring of the remaining graph, converges to 1. This ensures the connectivity with secure link in a sensor network with random key pre-distribution scheme.

This completes the proof of theorem.

Theorem 3: Let G_N be a sensor network, with N sensors and size of key pool, $K = \mid \mathcal{K} \mid = N \log N$ and size of key-ring an individual node X is $m = \mid k(X) \mid = c \log N$, where c is some positive constant. Then for sufficiently large N, and $c > \sqrt{2}$ network is connected almost surely.

Proof. Let $S \subset V$ be a cut in the graph. For showing that the network is connected, it is sufficient to prove that their is no cut subgraph $S \subset V$ in the graph G_N.

$$\begin{aligned} & \{\text{Network is not connected}\} \\ \subset & \{\exists S, \mid S \mid = s, k(S) \cap k(V \setminus S) = \phi\}. \end{aligned} \quad (5)$$

Then

$$\begin{aligned} & P[\text{Network is not connected}] \\ \leq & P[\exists S, \mid S \mid = s, k(S) \cap k(V \setminus S) = \phi]. \end{aligned} \quad (6)$$

Using (5), in above (6), we have

$$\sum_{N=1}^{\infty} P[\text{Network is not connected}] < \infty.$$

Now by the Borel-Cantelli's Lemma , we have

$$P[\text{Network is not connected} \qquad i.o.] = 0.$$

Hence, Network is connected almost surely.

IV. Probability of Connectivity by 1-hop Neighbors.

Let us consider a 1-hop neighbor: $\langle u, u_1, v \rangle$, where $u, v \in V$ are nodes not directly connected to each other and $u_1 \in V$ be an arbitrary intermediate nodes by which we are intersected to establish the secure connection between u and v.

Define an event Y such that $u, v \in V$ are 1-hop neighbors, i.e., $u, v \in V$ are not directly connected and having 1-hop connection through an arbitrary node $u_1 \in V$. Then

$$\begin{aligned} & P[Y] \\ = & P[u, v \in V \text{ are 1-hop neighbors.}] \\ = & 1 - P[u, v \in V \text{ are not directly connected}] \\ \cdot & P[u, v \in V \text{ are not connected through } u_1] \\ = & 1 - q_s \cdot P[u \text{ or } v \text{ are not connected } u_1] \\ = & 1 - q_s \\ \cdot & P[\{k(u) \cap k(u_1) = \phi\} \cup \{k(v) \cap k(u_1) = \phi\}] \\ = & 1 - q_s \\ \cdot & (1 - P[\{k(u) \cap k(u_1) \neq \phi\} \cap \{k(v) \cap k(u_1) \neq \phi\}]) \\ = & 1 - q_s(1 - p_s^2) = 1 - q_s^2(1 + p_s). \end{aligned} \quad (7)$$

(8)

Using (1) in (8), we get

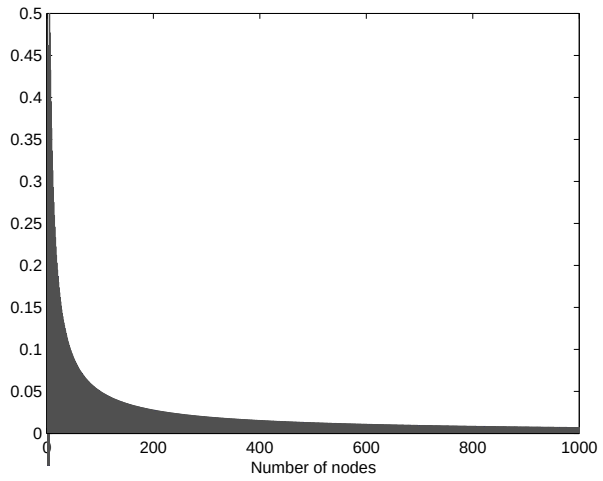
$$\begin{aligned} P[Y] & \leq 1 - \left(1 - \frac{m}{K-m}\right)^{2m} \left(1 + \frac{m^2}{K}\right) \\ & \leq 1 - \left(1 - \frac{2m^2}{K}\right) \left(1 + \frac{m^2}{K}\right) \\ & = 1 - \left(1 - \frac{m^2}{K} - \frac{2m^4}{K^2}\right) \\ & = \frac{m^2}{K} \left(2\frac{m^2}{K} + 1\right). \end{aligned} \quad (9)$$

Substituting $m = c \log N$ and $K = N \log N$, in (9). Then we have

$$P[Y] \leq \frac{c^2 \log N}{N} \left(2\frac{c^2 \log N}{N} + 1\right). \quad (10)$$

Now it is clear from the above expression that for sufficiently large N, the probability $P[Y]$ is converges to 0. Although the probability $P[Y]$ is not summable, we can not say anything about the almost sure convergence. From (10), it is clear that as number of the node increase in the network that number of 1-hop neighbor decrease in the network.

In the following figure we plot right hand side expression of (10) against the number of node in the network.



It is clear from the (10) that the probability that two arbitrarily chosen vertices $u, v \in V$ are 1-hop neighbors is always lying in the gray region show in the figure.

V. Appendix.

Proof of Lemma 1.

$$\begin{aligned} q_s &= \frac{\binom{K-m}{m}}{\binom{K}{m}} \\ &= \frac{(K-m)!(K-m)!}{(K-2m)!K!} \\ &= \left(\frac{(K-m)(K-m-1)\dots(K-2m+1)}{K(K-1)(K-2)\dots(K-m+1)}\right) \\ &= \left(1-\frac{m}{K}\right)\left(1-\frac{m}{K-1}\right)\dots\left(1-\frac{m}{K-m+1}\right) \end{aligned}$$

Then it is easy to show that

$$\left(1-\frac{m}{K-m}\right)^m \le q_s \le \left(1-\frac{m}{K}\right)^m \le \exp\left(-\frac{m^2}{K}\right) \tag{11}$$

From above we can derive the lower bound of q_s,

$$\begin{aligned} q_s &\ge \left(1-\frac{m}{K-m}\right)^m \\ &= 1-\frac{m^2}{K-m}+\frac{m^2(m-1)}{2}\left(\frac{m}{K-m}\right)^2\dots \\ &\ge 1-\frac{m^2}{K-m}. \end{aligned}$$

This implies

$$1-q_s \le \frac{m^2}{K-m}. \tag{12}$$

Now for the upper bound of q_s, we have

$$\begin{aligned} q_s &\le \left(1-\frac{m}{K}\right)^m \\ &= 1-\frac{m^2}{K}+\frac{m(m-1)}{2}\left(\frac{m}{K}\right)^2 \\ &\le 1-\frac{m^2}{K} \\ \Rightarrow 1-q_s &\ge \frac{m^2}{K}. \end{aligned} \tag{13}$$

From (12) and (13), we have

$$\frac{m^2}{K} \le p_s \le \frac{m^2}{K-m}.$$

This completes the proof.

VI. Conclusion.

In this paper, we have shown that we can almost surely achieve a secure connected network by using random key pre-distribution scheme as described in [6]. For achieving this aim in a N nodes network, the key pool size must be of order of $N \log N$ and the key ring size must be of order of $c \log N$, where $c > \sqrt{2}$ for sufficiently large N. Also we derive the upper bound for the probability that two arbitrarily chosen vertices have 1-hop connectivity.

References

[1] Das, A.K., "An identity-based random key pre-distribution scheme for direct key establishment to prevent attacks in wireless sensorn networks," *International Journal of Network sSecurity*, Vol.6, page 134-144, 2008.

[2] Castelluccia, C., and Spognardi, A., "RoK: A robust key pre-distribution protocal for multi-phase wireless sensor networks," *SecreComm*, France 2007.

[3] Chan, H., Perrig, A. and Song, D., "Random key predistribution scheme for sensor networks," In Proceedings of *IEEE symposium on security and privacy*,Oakland USA, page 197-213, 2003.

[4] Dai, T.T., and Hong, C.S., "Efficient ID-based threshold random key pre-distribution scheme for wireless sensor netwroks,"*IEICE Transaction on Communication*, Vol.E91-B, No.8, page 2602-2609, 2008.

[5] Du, W., Deng, J., Han, Y.S., and Varshney, P.K., "A pairwise key pre-distibution scheme for wireless sensor network," *CCS'03*, Washington, page 27-30, 2003.

[6] Eschenauer, L. and Gligor, V.D., "A key-managment scheme for distributed sensor networks," in Proceedings of the 9^{th} *ACM conference on Computer and communications security,*, page 41-47, 2002.

[7] Ghose, S.K., "On optimality of key pre-distribution schemes for distributed sensor networks," 3*rd European Workshop on Security and Privacy in Ad hoc and Sensor Networks ESAS*, 2006.

[8] Liu, D., and Ning, P., "LocationBased Pairwise Key Establishments for Static Sensor Networks," *SASN*, page 72-82, 2003.

[9] Pietro, R.Di and Mancini, L.V., "How to design a connected sensor networks that are provable secure," in Proceedings of the *SecureComm, the2nd IEEE/CreatNet International conference on security and privacy in communications networks*, Baltimore 2006 .

[10] Yagan, O. and Makowski, A.M., "On the random graph induced by a random key predistribution scheme under full visibility," in Proceedings of the *ISIT 2008.*

Bhupendra Gupta PhD Statistics, from Department of Mathematics and Statistics, IIT Kanpur in 2008. He was born in Meerut, U.P., India on 28th Nov. 1976. Presently, he is working as assistant professor in Indian Institute of Information Technology, Design & Manufacturing Jabalpur, MP, India. Dr. Gupta's area of interest is random networks and their application in various areas like sensor networks etc. His main research interests include communication networks and performance analysis. His current research has concentrated on random networks with applications in Network security, wireless and sensor networks.

Subir Singh Lamba PhD Mathematics, from Department of Mathematics and Statistics, IIT Kanpur in 2006. He was born in Lucknow, U.P., India on 12th Jan. 1971. Presently, he is working as assistant professor in Indian Institute of Information Technology, Design & Manufacturing Jabalpur, MP, India. His area of interest are parallel computing, Spectral Method and Numerical Solution to PDE's.