Key Pre-distribution in Wireless Sensor Networks Using Key Groups

Chih-Shiang Chang and Tien-Ruey Hsiang
Department of Computer Science and Information Engineering
National Taiwan University of Science and Technology
{M9315031, trhsiang}@mail.ntust.edu.tw

ABSTRACT

Prior results on the key pre-distribution in wireless sensor networks often employ the deployment knowledge of the sensor nodes to construct a better communication graph. This paper proposes the use of key groups in random key pre-distribution. Three different approaches are discussed for key selection and experiments are conducted to compare the effectiveness with prior proposed schemes. Through the simulation results, we see that with the help of key groups, a smaller number of keys can be preloaded in the sensor nodes to achieve the same level of robustness of the network.

Index Terms—key management, sensor networks, random graphs, probabilistic key sharing

1: INTRODUCTIONS

Distributed sensor networks consist of a large number of simple wireless devices. Each device, called a sensor node, has limited power and low building cost, which restrict the node to have only limited communication and computation capabilities and limited memory capacity. There is no fixed structure in a sensor network since its topology changes over time. Each node finds the nearest node in its communication range to achieve long-range data transmission.

There are many applications for sensor networks, including military applications such as mine sweeping, environment monitoring, patient monitoring and tracking, smart environments, etc. When sensor networks are deployed in hostile environments, secure communication becomes extremely important. Key pre-distribution is a widely used method to establish a secure system in sensor networks. In order to realize a key pre-distribution scheme, sensor nodes must use a key pool with a large number of unique keys. Before the deployment, every node selects a number of keys from the key pool and stores these keys in the memory. After sensor nodes are deployed, when a certain node wants to communicate with its neighboring node, they exchange the information of keys. If the shared keys satisfy the minimum requirement, a secure communication link can be established between two nodes.

2: RELATED WORK

Random graphs are often used in studies of the key pre-distribution problem. A random graph is a graph where the probability that an edge exists between any two vertices is p. Let $G(n,p)$ be a random graph with n vertices, then G has no edge if p=0 and G is fully connected if p=1. A vertex in $G(n,p)$ has expected degree $d = p*(n-1)$. Erdos and Renyi [9] showed that the probability of a random graph being connected is

$$Pc = \lim_{n\to\infty} \Pr[\ G(n,p)\ is\ \ connected\ \ ] = e^{-e^{-c}},$$

where

$$p = \frac{\ln(n)}{n} + \frac{c}{n} \quad \text{and } c \text{ is real constant.}$$

Eschenauer and Gligor [4] proposed the random key-chain based key pre-distribution scheme. It includes three phases, which are the key pre-distribution phase, the shared-key discovery phase, and the path-key establishment phase. In the key pre-distribution phase, the key pool generates a large number of keys, the number of keys is between $2^{17} \sim 2^{20}$. After establishing the key pool, each node randomly selects k keys from the key pool. In the share-key discovery phase, any two nodes try to find its neighbor to establish a secure link. Two nodes can conduct secure communication if they share a common key. Finally in the path-key establishment phase, each node tries to establish a path-key with other nodes that are in the communication range but do not share a common key.

Chan, Perrig and Song [5] proposed the q-composite scheme. For any pair of nodes sharing at least q common keys, a secure communication link can be established between them. This scheme enhances the previous scheme. It is more difficult to break because the attacker must collect more keys to successfully attack the network.

There are key pre-distribution schemes that use the deployment knowledge of sensor nodes. Liu, Ning and Du [2] studied group-based key pre-distribution in sensor networks. Their scheme divides nodes into several groups. Nodes that are in the same group are more likely to be neighbors. It defines two kinds of groups, that is, the in-group instances and the cross-group instances. Each node belongs to one in-group instance and one cross-group instance. Based on this model, their paper developed a novel group-based key pre-distribution scheme, which can be combined with many existing key pre-distribution techniques.

Zhou, Ni and Ravishankar [7] proposed a two-phase scheme consisting of two phases, the intra-group key pre-distributing phase and the inter-group key pre-distributing phase. In the intra-group phase, it preloads each pair of sensors from the same group with a unique pairwise key. In the inter-group phase, it selects any two nodes as an agent between any two groups. An agent shares a pairwise key so that neighboring sensors from any two groups can establish path keys using an agent as an intermediary.

3: KEY PRE-DISTRIBUTION SCHEMES AND EXPERIMENTS

Our schemes consist of two phases. When the key pool is generated, it is divided into several groups. This step is called the *key grouping phase*. If the key pool has P keys and it is divided into g groups, each group has $p' = \frac{P}{g}$ keys. After initializing the key groups, the *distributing phase* randomly select keys to be preloaded in sensor nodes without using the deployment knowledge.

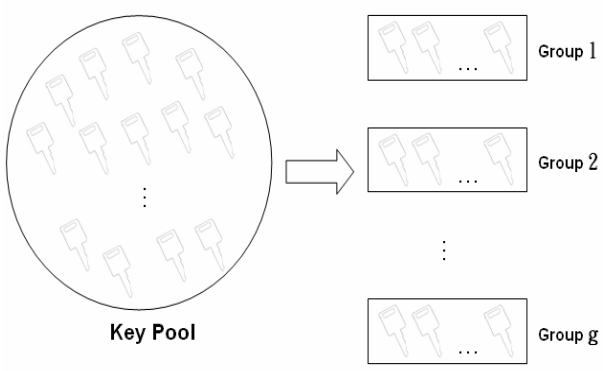

Fig. 1. Key grouping phase.

We discuss three approaches for key selection in the distributing phase, the "randomly select 1 group", the "relaxed randomly select 1 group" and the "randomly select 2 groups ".

3.1: Randomly Select-One-Group. After finishing the key grouping phase, each node randomly selects one out of the g key groups, then selects k keys from this group. We call this method the "randomly select-one-group" scheme (RSO).

The probability that two nodes share at least one key is $P_g = 1 - Pr$[two nodes do not share any key]. There are two cases that two nodes do not share the same keys. One is that two nodes select the same group but they have no shared keys, the other is that two nodes select different groups. Therefore, we have

$$P_g = 1 - \frac{\binom{g}{1} * \binom{p'}{k} * \binom{p'-k}{k} + \binom{g}{2} * \binom{p'}{k}^2 * 2}{\binom{g}{2} * \binom{p'}{k}^2}.$$

We use the above formula to conduct experiments and compare them with [2], which we refer to as "the basic scheme". We assume that the key pool size P is 10,000, the number of nodes is 100, and each node selects 200 keys.

Fig. 2 illustrates that the probability that any two nodes can communicate through a path. When P is 10,000 and the group number g is 2, we see that our scheme's probability is less than the basic scheme proposed in [4]. However, when P is 100,000 in the basic scheme and the group number g is 5, our scheme connects better than the basic scheme when nodes select less than 150 keys.

In Fig. 3, we show the threshold value of the number of captured nodes for an attacker to break the network, that is, the minimum number of compromised nodes that an attacker needs to communicate with any single node. When the group numbers exceed 5, the attacker needs to capture more nodes in order to break the network. Our scheme shows better effectiveness than the basic scheme.

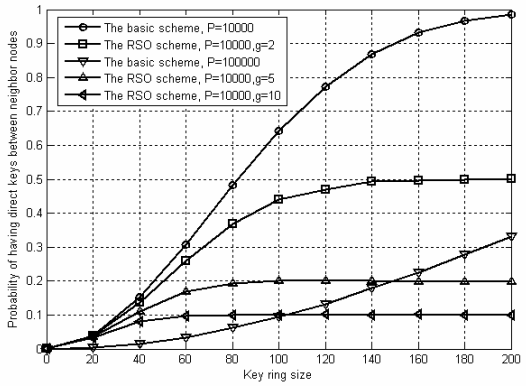

Fig. 2. The probability of sharing at least one key when each of a pair of nodes chooses k keys from a pool of size P. The RSO scheme compares with the basic scheme.

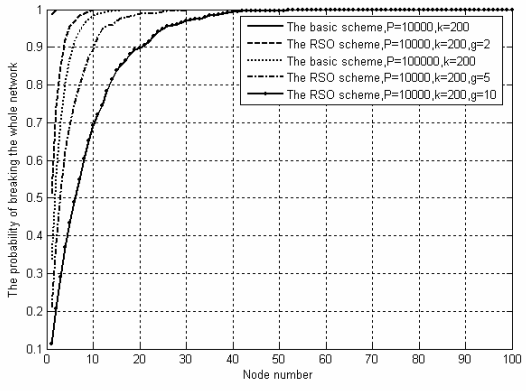

Fig. 3. The probability of an attacker breaking the network by capturing nodes in RSO.

From Fig. 1 and Fig. 2, we know that both P and g play important roles in RSO. To better understand the impact of these factors, we vary the number of nodes and the key ring size in the experiments. The probability of the attacker using exactly one captured node to break the network as the number of nodes increases is stable, which is shown in Fig. 4. Fig. 5 has a similar result as the key number in a node increases. Regardless of the

groups' and the key pool's sizes, the probability remains constant. Thus, in order to prevent the attack, we can adjust the size of the key pool and the number of groups. In the rest of this paper, we will use these two factors to conduct experiments.

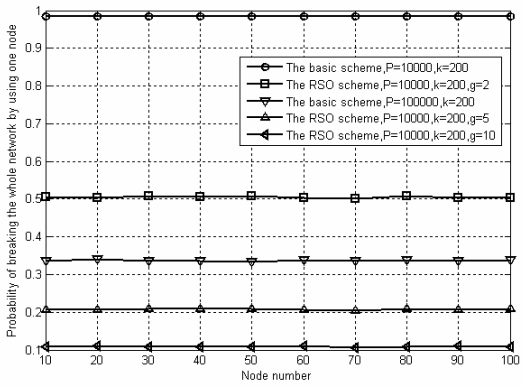

Fig. 4. The probability of an attacker breaking the network by capturing any one node as the node number increases in RSO.

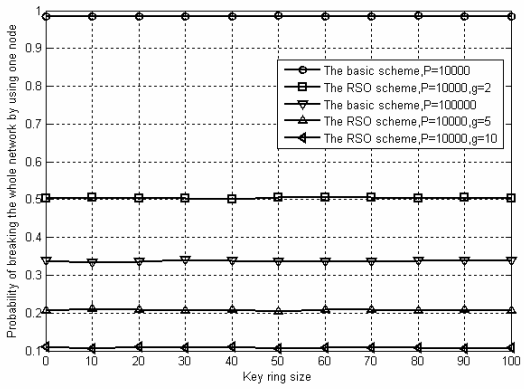

Fig. 5. The probability of an attacker breaking the whole network by capturing any one node as the key ring size increases.

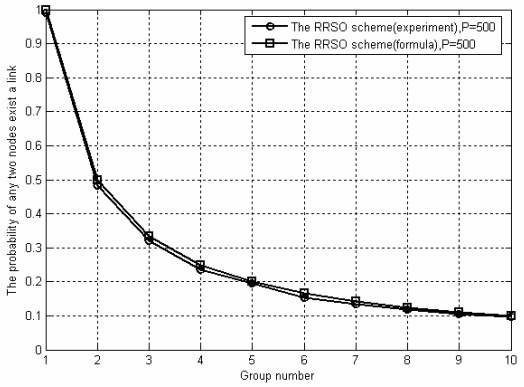

Fig. 6. The probability of two nodes sharing at least one key.

3.2: Relaxed Randomly Select-One-Group. After finishing the key grouping phase, each node randomly selects one out of the g groups, then selects at most p' keys from this group. Under this setting, the size of each node's key ring is different. We call this scheme the "relaxed randomly select-one-group" scheme (RRSO).

Since each node can select at most p' keys of a group, there are $[g * \sum_{i=1}^{p'} (\binom{p'}{i})]^2$ possible cases. There are two cases that two nodes do not share any key. One case is that two nodes select the same group; the other is that the two nodes select different groups. Therefore, the probability P_g is

$$P_g = 1 - \frac{2*\binom{g}{2}*[\sum_{i=1}^{p'}(\binom{p'}{i})]^2 + \binom{g}{1}*\{\sum_{i=1}^{p'}[(\binom{p'}{i})*\sum_{j=1}^{p'-i}(\binom{p'-i}{j})]\}}{[g*\sum_{i=1}^{p'}(\binom{p'}{i})]^2}$$

We conduct experiments and compare them with the basic scheme. We assume that key pool size P is 500 and the number of nodes equals100.

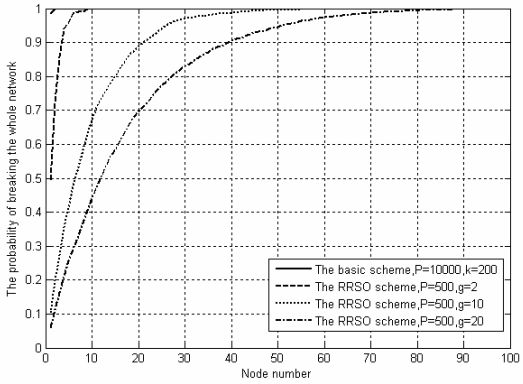

Fig. 7. The probability of an attacker breaking the whole network by capturing nodes.

Fig. 6 shows the probability P_g with the above equation and the experiment according to the group number. Fig. 6 also validates our formula. The probability is 0.5 and 0.1 when group is 2 and 10, as shown in Fig. 6. Note that this result corresponds with Fig. 2. Also, in the resilience experiment, the result we see in Fig. 7 is similar to what is shown in Fig. 3. From Fig. 6 and Fig. 7, we get the same effectiveness with the RSO scheme with smaller key pool size.

3.3: Randomly Select-Two-Groups. This scheme combines the ideas used in the RSO and the q-composite schemes. After finishing the key grouping phase, each node randomly picks two out of g groups, then selects k keys arbitrarily from these two groups. A pair of nodes must share q keys to establish a communication link between them. We call this scheme the "randomly select-two-group" scheme (RST).

We note that the probability of two nodes sharing at least q key is $P_g = 1 - \sum_{i=0}^{q-1} \Pr$ (share i keys). Since each node selects two groups, there are three cases that two nodes do not share any keys. The first case is that the two nodes select different groups and its probability p_1 is $p_1 = \frac{\binom{g}{2}*\binom{g-2}{2}}{\binom{g}{2}^2}$. The second case is that exactly one group is shared by the two nodes. The number of possible cases that two nodes share q keys in one group is $\binom{p'}{q}$. Because each of the two nodes selects q keys,

they need *k-q* keys to select. So the number of possible cases is $\binom{p'-q}{2*(k-q)}$. These keys must be distributed equally to two nodes, which has $\binom{2*(k-q)}{k-q}$ cases. The number of possible cases of group selection is $\binom{g}{3}*3$. We get

$$p_2 = \frac{\binom{p'}{q}*\binom{p'-q}{2*(k-q)}*\binom{2*(k-q)}{k-q}}{\binom{p'}{k}^2} * \frac{\binom{g}{3}}{\binom{g}{2}^2} * 3$$

The remaining case is that the two nodes select the identical set of two groups. If a pair of nodes sharing *i* keys in the first group, then they must share *q-i* keys in the other group. The number of possible cases of group selection is $\binom{g}{2}$, we get

$$p_3 = \frac{\binom{g}{2}}{\binom{g}{2}^2} * [\frac{\binom{p'}{i}*\binom{p'-i}{2*(k-i)}*\binom{2*(k-i)}{k-i}}{\binom{p'}{k}^2} * \frac{\binom{p'}{j}*\binom{p'-j}{2*(k-j)}*\binom{2*(k-j)}{k-j}}{\binom{p'}{k}^2}],$$

where *i=1~q-1* and *j=q-1~1*.

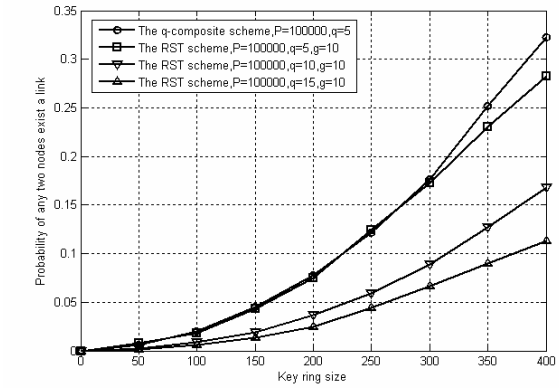

Fig. 8. The probability of two nodes sharing at least *q* keys. The RST scheme compares with the q-composite scheme.

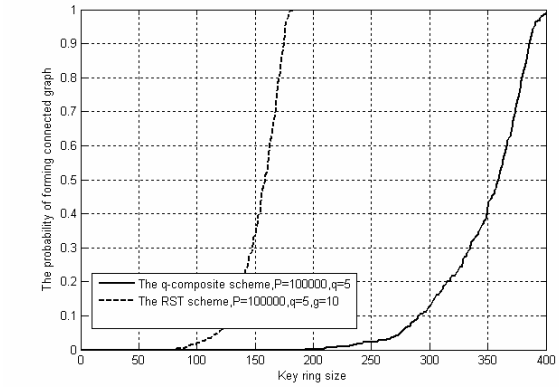

Fig. 9. The probability of forming a connected communication graph.

To compare the RST scheme with the *q*-composite scheme, we set the key pool *P* to be 100,000, the number of nodes be 100, and the key ring size be 400. Fig. 8 illustrates the probability of two nodes being able to communicate directly. When *q* is 5, the *q*-composite and RST has roughly the same probability until key ring is 300. However, the probability of communication graphs being connected is quite different, as shown in Fig. 9. We see that RST has a higher probability of connection than the *q*-composite. The difference between Fig. 8 and Fig. 9 is due to each node can select two groups in RST scheme. If any two nodes select at least one identical group, then the probability of these two nodes having a link must be very high. To be clear, we separate the situation into two cases as shown in Fig. 10. The first selected group of Node A maybe equal to either the first selected group of Node B or the second selected group of Node B. The simulation results of the two cases are in Fig.11 and Fig.12, respectively.

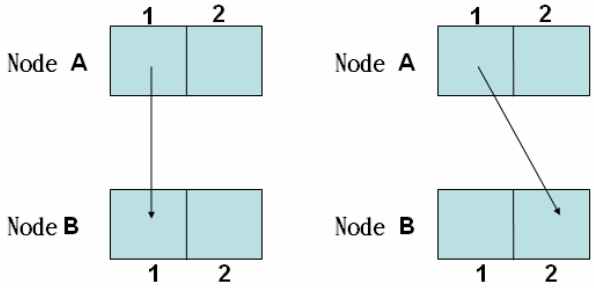

Fig. 10. Two cases of two nodes selecting at least one identical group.

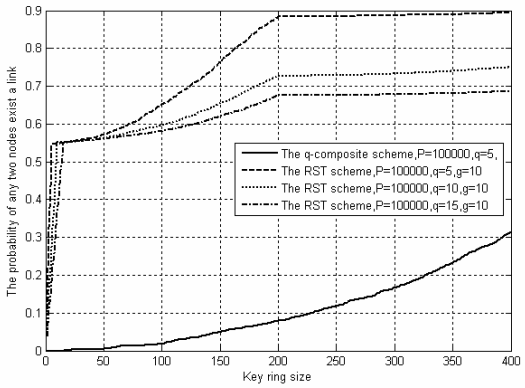

Fig. 11. The probability of sharing at least *q* keys in case 1 of Fig. 10.

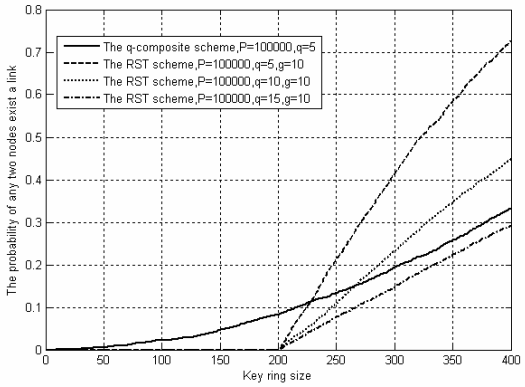

Fig. 12. The probability of sharing at least *q* keys in case 2 of Fig. 10.

The probability of any two nodes sharing at least *q* keys in RST scheme is higher than what is obtained by the *q*-composite scheme. Case 1 and case 2 occupy about 38% when each node selects groups. Along with Fig. 11 and Fig. 12, we get that the *q*-composite scheme has higher probability of any two nodes share *q* keys

than RST and RST has higher probability of connected graph than the *q*-composite scheme.

Fig. 13 shows the result of the resilience experiment. When *q* is 5, an attacker under RST needs more nodes to break the network, which implies that RST scheme is indeed stronger against the attacks, and the probability of connectivity is higher than the *q*-composite scheme when nodes select the same key numbers.

Fig. 14 compares the probability of a direct key between two non-compromised sensor nodes being compromised for the group-based scheme in [2] and the RST scheme, where the total number of nodes is 5000. When *q* is 5 and there are 10 groups, the probability of the RST scheme is less than the *q*-composite scheme; however, it is higher than the group-based scheme. Because the RST curve stays close to group-based EG curve and the RST scheme is simple relative to the group-based scheme, the RST scheme is reliable.

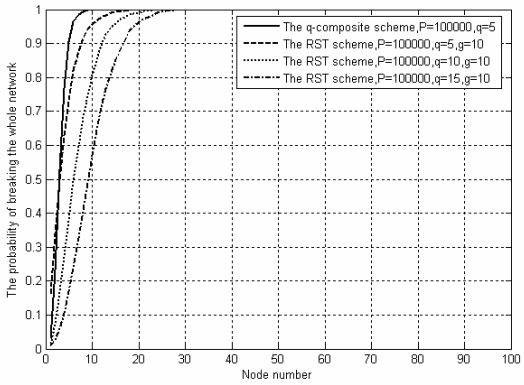

Fig. 13. The probability of an attacker breaking the whole network by capturing nodes.

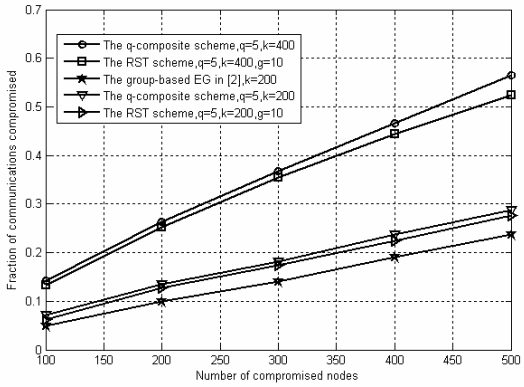

Fig. 14. The number of compromised nodes vs. the fraction of compromised link. The RST scheme compares with the q-composite scheme and group-based scheme in [2].

4: CONCLUSION

In this paper, we improve the basic scheme in [4] and discuss three extensions. Our main idea is to divide the key pool into several groups, and then pre-distribute keys to each sensors randomly in terms of both the key groups and the keys in each group. Our schemes have some advantages over the basic scheme and the *q*-composite scheme and the computation overhead is low between any pair of nodes.

The RSO scheme only uses 10% key numbers compared to the basic scheme and it is harder to break. The RRSO scheme only uses 500 key numbers in the key pool, to achieve better effectiveness than the RSO scheme. Finally, we combine the RSO and the *q*-composite schemes to form the RST scheme. The probability of forming a connected communication graph in RST is better than the basic scheme and *q*-composite and RST is stronger against attacks.

ACKNOWLEDGMENT

This work was supported in part by the National Science Council under the Grants NSC95-3114-P-001-001-T02, and by the Taiwan Information Security Center(TWISC) under the Grants NSC 94-3114-P-011-001, NSC 94-3114-P-001-001-Y, NSC94-3114-P-001-002-Y and NSC94-3114-P-001-003-Y.

REFERENCES

[1] C. Blundo, A. De Santis, A. Herzberg, S. Kutten, U. Vaccaro and M. Yung, *Perfectly Secure Key Distribution for Dynamic Conferences*. In Advances in Cryptology --- CRYPTO '92, pages 471--486. Springer-Verlag, Berlin, 1993.

[2] D. Liu, P. Ning and W. Du. *Group-based key pre-distribution in wireless sensor networks*. In Proceedings of the 4th ACM Workshop on Wireless Security (Cologne, Germany, September 02 - 02, 2005). WiSe '05, pages 11--20, ACM Press, 2005.

[3] W. Du, J. Deng, Y.S. Han. P.K.Varshney. *A Pairwise Key Pre-distribution Scheme for Wireless Sensor Networks*. In Proceedings of the 10th ACM Conference on Computer and Communication Security, pages 42--51, 2003.

[4] 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, pages 41--47. ACM Press, 2002.

[5] H. Chan, A. Perrig, and D. Song. *Random key predistribution schemes for sensor networks*. In IEEE Symposium on Security and Privacy, pages 197--213, May 2003.

[6] J. Hwang, Y. Kim. *Revisiting Random key Predistribution Schemes for Wireless Sensor Networks.* In SASN '04: Proceedings of the 2nd ACM workshop on Security of ad hoc and sensor networks, pages 43--52. ACM Press, 2004.

[7] L Zhou, J Ni, CV Ravishankar. *Efficient key establishment for group-based wireless sensor deployments.* In Proceedings of the 4th ACM Workshop on Wireless Security (Cologne, Germany, September 02 - 02, 2005). WiSe '05, pages 1--10. ACM Press, 2005.

[8] D. Liu and P. Ning. *Establishing pairwise keys in distributed sensor networks*. In CCS '03: Proceedings of the 10th ACM conference on Computer and communications security, pages 52--61. ACM Press, 2003.

[9] P. Erdos and A. Renyi. *On the evolution of random graph.* Institute of Mathematics Hungarian Academy of Sciences, pages 17--61, 1959.

[10] R. Blom, *An Optimal Class of Symmetric Key Generation Systems*. Advances in Cryptology: Proc. of Eurocrypt 84, Lecture Notes in Computer Science, 209, Springer-Verlag, Berlin, pp. 335--338, 1984.

[11] S. Camtepe, B Yener . *Key Distribution Mechanisms for Wireless Sensor Networks: a Survey.* Technical Report, Rensselaer Polytechnic Institute, 2005.