

A Survey about Key Pre-distribution Scheme in Wireless Sensor Networks

Amol Abhiman Magar
Mtech-Comp. sci. and Technology.
Email: amolmagar@outlook.com
Contact no: 0855 195 0789

Abstract— Wireless sensor network (WSN) has a wide range of applications in military as well as in civilian services. As wireless sensor networks continue to grow, so does the need for effective security mechanisms. Because sensor networks may interact with sensitive data and/or operate in hostile unattended environments and also nearly all aspects of wireless sensor network defenses relying on solid encryption and mainly key predistribution is a challenging task in sensor networks. Because the neighbor of a node after the deployment of sensors is unknown. An attacker can easily obtain a large number of keys by capturing a small fraction of nodes, and hence, can gain control of the network by deploying a replicated nodes or packets preloaded with some compromised keys. For secure communication, neighbors must possess a secret common key or there must exist a key-path among these nodes. In this paper I have discussed in brief about various key pre-distribution schemes for homogeneous sensor networks and I had analyzed merits and demerits for each of them. Among various schemes a suitable scheme can be chosen based on the requirement and the resource availability of the sensors.

Keywords— Wireless Sensor Networks, Key Pre-Distribution, Resiliency, BIBD, Mobile Polynomial Pool, Grid Based Key, Hybrid Key, Pair Wise Key Establishment.

Introduction

Wireless sensor network (WSN) has a wide range of applications in military as well as in civilian services. Sensor nodes are deployed in a battlefield to detect enemy intrusion. They are used to measure various environmental variables such as temperature, heat, sound, pressure, magnetic and seismic fields, etc. of a region. It has several applications in industry such as machine health monitoring, waste water monitoring etc. As the sensor nodes are used in various applications, secure communication between the sensor nodes is needed in order to keep the information secret. For secure communication between two sensor nodes a secret key is needed and cryptographic key management is a challenging task in sensor networks. Sensor nodes are constrained in resources, such as they have low processing power, less memory capacity, limited battery life. Apart from this wireless nature of the network, unknown topology of the network, higher risk of node capture and lack of fixed infrastructure makes the key management more challenging in WSN. Use of any cryptographic algorithm must take into account the resource availability at each node. They should be easy in computation and occupy less storage space. Use of symmetric key cryptography means each node should maintain $(N-1)$ keys for N number of nodes in the network. For a large value of N , a substantial memory space is wasted in storing the keys, and hence is not memory efficient. Use of public key cryptosystem needs a huge computational power and sensors have low processing power. Hence public key cryptosystem is not an efficient key management technique in WSN. Key pre-distribution scheme is regarded as a promising key management in sensor network. In key pre-distribution scheme, each sensor is assigned a set of keys from a pool of keys before deployment such that after deployment, two nodes who are in the communication range of each other will share at least one key between them with higher probability, so that a secure communication can be established between them.

TERMS AND DEFINITIONS

In this section I briefly discuss some of the related terminologies and definitions for the sake of completeness. A set system or design [26] is a pair (X, A) , where A is a set of subsets of X , called blocks. The elements of X are called varieties or elements. A Balanced Incomplete Block Design BIBD $(v; b; r; k; \lambda)$, is a design which satisfies the following conditions:

- 1) $|X|=v, |A|=b$.
- 2) Each subset in A contains exactly k elements,
- 3) Each variety in X occurs in r blocks,
- 4) Each pair of varieties in X is contained in exactly λ blocks in A .

When $v = b$, the BIBD is called a symmetric BIBD (SBIBD) and denoted by $SB[v, k, \lambda]$.

An association scheme with m associate classes on the set X is a family of m symmetric anti-reflexive binary relations on X such that:

- 1) Any two distinct elements of X are i -th associates for exactly one value of i , where $1 \leq i \leq m$.
- 2) Each element of X has n_i i -th associates, $1 \leq i \leq m$.

3) for each i , $1 \leq i \leq m$, if x and y are i -th associates, then there are p^{ij} elements of X which are both j -th associates of x and i -th associates of y . The numbers v , n_i ($1 \leq i \leq m$) and p^{ij} ($1 \leq i, j; i, j \leq m$) are called the parameters of the association scheme.

A partially balanced incomplete block design with m associate classes, denoted by PBIBD (m) is a design on a v -set X , with b blocks each of size k and with each element of X being repeated r times, such that if there is an association scheme with m classes defined on X where, two elements x and y are i -th ($1 \leq i \leq m$) associates, then they occur together in λ_i blocks. We denote such a design by PB[$k; \lambda_1; \lambda_2; \dots; \lambda_m; v$].

Let X be a set of varieties such that

$$X = \bigcup_{i=1}^m G_i, |G_i| = n \text{ for } 1 \leq i \leq m, G_i \cap G_j = \emptyset \text{ for } i \neq j$$

The G_i 's are called groups and an association scheme defined on X is said to be group divisible if the varieties in the same group are first associates and those in different groups are second associates.

A transversal design TD (k, λ, r), with k groups of size r and index λ , is a triple (X, G, A) where,

- 1) X is a set of kr elements (varieties).
- 2) $G = (G_1, G_2, \dots, G_k)$ is a family of k sets (each of size r) which form a partition of X .
- 3) A is a family of k -sets (or blocks) of varieties such that each k -set in A intersects each group G_i in precisely one variety, and any pair of varieties which belong to different groups occur together in precisely λ blocks in A .

KEY PRE-DISTRIBUTION SCHEMES

All the key pre-distribution schemes can be divided into three according to the way of choosing keys for each node from the key pool. They are :

- 1) Probabilistic: Keys are drawn randomly and placed into the sensors.
- 2) Deterministic: Keys are drawn based on some definite pattern.
- 3) Hybrid: Makes use of both the above techniques.

To discuss about the schemes in a better way we have divided them into some parts and we have discussed below about each part in respective subsections.

A. Basic schemes

First I'll discuss about two basic schemes which though were not meant for WSN, have been used in context of WSN. Those two schemes are Blom's scheme and Blundo et al's scheme.

Blom [1] proposed a key pre-distribution scheme that allows any two nodes of a group to find a pair wise key. The security parameter of the scheme is c , i.e., as long as no more than c nodes are compromised, and the network is perfectly secure. They have used one public matrix and one secret symmetric matrix to construct this scheme. Each node will have the share of those matrix such that any two nodes can calculate a common key between them without knowing each other's secret matrix share. The problem with this scheme is that if more than c numbers of nodes are compromised, the whole network will be compromised.

In the scheme proposed by Blundo, Santis, Herzberg, Kutten, Vaccaro, Yung [2], they used a symmetric bivariate polynomial over some finite field $GF(q)$. Symmetric bivariate polynomial is a polynomial $P(x, y) \in GF(q)[x, y]$ with the property that $P(i, j) = P(j, i)$ for all $i, j \in GF(q)$. A node with ID U_i stores a share in P , which is an univariate polynomial $f_i(y) = P(i, y)$. In order to communicate with node U_j , it computes the common key $K_{ij} = f_i(j) = f_j(i)$; this process enables any two nodes to share a common key. If P has degree t , then each share consists of a degree t univariate polynomial; each node must then store the $t + 1$ coefficients of this polynomial. So, each node requires space for storing $t + 1$ keys. If an adversary captures s nodes, where $s \leq t$, then it cannot get any information about keys established between uncompromised nodes. However, if it captures $t + 1$ or more nodes then all the keys of the network can be captured.

Eschenauer and Gligor first proposed a random key predistribution scheme [12] for WSN. They divided the key predistribution mechanism into three steps: key pre-distribution, shared-key discovery and path-key establishment. In this approach, a key ring for a node containing some fixed number of keys are chosen randomly without replacement from a key pool of large number of keys. Each node is assigned a key ring. The key identifiers of a key ring and corresponding sensor identifiers are stored in a trusted controller

node. Now a shared key may not exist between two nodes. In that case, if there exists a path of nodes sharing keys pairwise between those two nodes, they may communicate via that path. They have also shown that for a network of 10000 nodes, a key ring containing 250 keys is enough for almost full connectivity. When sensor nodes are compromised, key revocation is needed. For this a controller node broadcasts a revocation message containing the list of identifiers of keys which have been compromised and all the nodes after getting the message removes the compromised keys from the key ring. The main advantages of this scheme are that the scheme is flexible, scalable, efficient and easy to implement. However, the main disadvantages are that it cannot be used in regions which are prone to massive node capture attack.

Chan Perrig and Song [8] modified Eschenauer and Gligor scheme. According to their q-composite scheme two nodes must share at least q number of keys to have a secure path between them. The path key will be formed by the hash of all the common keys. Though for small number of node capture, resiliency was improved, the resiliency was affected drastically as number of captured nodes increases.

B. Random pair wise scheme

In the random pairwise scheme, proposed by Chan, Perrig and Song [8], they have proposed that in a network of size N and minimum connection probability of two nodes is p, each node will store k number of keys where $k = N * p$. The key pre-distribution, shared key discovery and path key establishment is done as in [12]. Node revocation for compromised nodes are done by voting of all the nodes in the network with a suitable threshold parameter. But the disadvantage of this scheme is that it is not scalable and choosing the threshold value for node revocation is very important as it can lead to other problems.

The pairwise key scheme of Liu and Ning [15] is based on the polynomial pool based key pre-distribution by Blundo et al [2]. They have shown the calculation for the probability that two nodes share a common key. They have also shown the probability that a key is compromised. Later it was extended in [16] where they modified the scheme into a hypercube based key pre-distribution. Zhu, Xu, Setia and Jajodia [25] also proposed a random pairwise scheme based on probabilistic key sharing where two nodes can establish shared keys without the help of an online KDC and only knowing each other's key id. Communication overhead in this scheme is very low. But if any node in the path is compromised then the key establishment process has to be restarted.

C. Grid-based key pre-distribution schemes

Chan and Perrig was the first to propose a grid based key pre-distribution scheme where they place all the nodes of a network in a square grid. The scheme was named as PIKE scheme [7]. In that scheme, each node will have a secret pairwise key with the nodes which lie in the same row or same column. So for a network of size N, each node has to store $2(\sqrt{N} - 1)$ number of keys. If two nodes do not have any shared key, they will have exactly two intermediate nodes having shared key with both the nodes. Here any node can act as an intermediary. Hence, it reduces the battery drainage of the nodes near base station who have to serve as intermediary most of the time in other schemes. But the main disadvantage of this scheme is that it has high communication overhead. Because large number of key pairs will not have common key between them, path-key establishment will be very much time consuming.

In [20], Kalindi et. al. modified the PIKE scheme. They placed the nodes as well as the keys in a grid and divide the grid into some sub-grids. A node will have all the keys in its key chain which lie in its same row or column and which are in its same or neighboring sub-grids. Key needed to store in each node can be much less than [7] if number of subgrids are more. It will increase the resiliency but decrease the connectivity. The reverse will happen if number of sub-grids is lesser. Nodes belonging to the same sub-grid and in same row or same column share more keys. But they are not allowed to use all the common keys because capturing of one node of a row or column will reveal all the keys of that row and column.

Sadi, Kim and Park [21] proposed another grid based random scheme based on bivariate polynomials. In this scheme, they will first arrange the nodes into a $m \times m$ square grid. After that some $2m\omega$ bivariate polynomials will be generated and they will be divided into some group such that each row and each column will be assigned one group of polynomials. A node then will select some 2τ number of polynomials from its row polynomial group and column polynomial group. If two nodes are in same row or in same column, they use a challenge response protocol to find whether they are sharing a common polynomial. If they a shared polynomial, they can setup a shared key. Otherwise they will have to go for path key establishment and they will have to find two other intermediate nodes such that a path can be established. In this case also the communication overhead is high.

D. Group based key pre-distribution

Liu, Ning and Du observed that sensor nodes in the same group are usually close to each other and they proposed a group based key pre-distribution scheme without using deployment knowledge [17], [16]. They divide the nodes of a network into groups and then form cross groups taking exactly one sensor node from each group such that there will not be any common node between any two cross groups. They presented two instantiations of pre-distribution. In the first one, hash function was used. Two nodes will share a common key if they are in same group or in same cross group. If the number nodes in the network are N and they are divided into n groups each containing m nodes, $N = n \times m$ and each node need to store $(m+n) / 2$ keys. In the second method, they used symmetric bivariate polynomials and assign a unique polynomial to each group and cross group. Every node will have share of the polynomials corresponding to their groups and cross groups. The advantages of this scheme are that it does not do not use deployment knowledge and give resiliency and connectivity similar to the deployment knowledge based schemes. The polynomial based schemes can be made scalable. The framework can be used to improve any existing predistribution schemes. The disadvantages of this scheme is that the probability of secure communication between cross-group neighbors is very less. The scheme is not suitable for networks which have small group size.

To overcome the problems of Liu et al's scheme [16], Martin Paterson and Stinson [19] proposed a group based design using resolvable transversal designs. To increase the cross group connectivity, they proposed that each node is contained in m cross groups rather than one. Though some additional storage is required. They did not give any algorithm for the construction of such designs.

E. Key pre-distribution using combinatorial structures

In the schemes which use combinatorial structures, one of their greatest advantages is that almost all of them have efficient shared key discovery algorithm with which easily two nodes can find their common key. Camtepe and Yener were the first to use combinatorial structures in key predistribution [4], [3]. They have used projective planes and generalized quadrangles. A finite projective plane $PG(2, q)$ (where q is a prime power) is same as the symmetric BIBD, $BIBD(q^2+q+1, q^2+q+1, q+1, q+1, 1)$. So, q^2+q+1 number of nodes can be accommodated in the network each node having $q + 1$ number of keys. It ensures 100% connectivity. But the resiliency was very poor. So they used generalized quadrangles, $GQ(s,t)$ where s and t are the two parameters of GQ . Three designs were used : $GQ(q,q)$ was constructed from $PG(4,q)$, $GQ(q,q^2)$ was constructed from $PG(5,q)$, $GQ(q^2, q^3)$ was constructed from $PG(4, q^2)$. Camtepe and Yener have mapped these GQs in key pre-distribution [4], [3] like this:

$v =$ number of keys $= (s + 1)(t + 1)$, $b =$ number of nodes $= (t + 1)(st + 1)$, $r =$ number of keys in each node $= (s + 1)$, and $k =$ key chains that a key is in $= (t + 1)$ for all the three GQs, these parameters are given in Table - 1. Here q is taken as any prime or prime power.

TABLE I

VARIOUS GENERALIZED QUADRANGLES USED BY CAMTEPE YENER AND THEIR DIFFERENT PARAMETERS

Design	s	t	v	b	k	r
$GQ(q, q)$	q	q	$(q + 1)(q^2 + 1)$	$(q + 1)(q^2 + 1)$	$q + 1$	$q + 1$
$GQ(q, q^2)$	q	q^2	$(q + 1)(q^3 + 1)$	$(q^2 + 1)(q^3 + 1)$	$q + 1$	$q^2 + 1$
$GQ(q^2, q^3)$	q^2	q^3	$(q^2 + 1)(q^3 + 1)$	$(q^3 + 1)(q^3 + 1)$	$q^2 + 1$	$q^3 + 1$

Probability that two node will share a common key in these GQs are $t(s+1) / [(t+1)(st+1)]$. Though GQs do not give 100% connection probability, resiliency is much better than projective planes.

Lee and Stinson [13] formalized the definitions of key predistribution schemes using set systems. They introduced the idea of common intersection designs [23]. They used block graphs for sensors and according to them; every pair of nodes can be connected by maximum of 2-hop path. They have shown that (v, b, r, k) -1 design or the (v, b, r, k) configuration have regular block graphs with vertex degrees maximized. So, connectivity will be largest in this case. So, they have used (v, b, r, k) configuration. In a (v, b, r, k) configuration having $b-1 = k(r-1)$, all the nodes are connected to each other and it's same as projective planes. But for large network, the keychain

in each node will be large. So, they introduced μ - common intersection design. In that if two node's key chain, A_i and A_j are disjoint, then there will be at least μ number of nodes, A_h who has common keys with both A_i and A_j . So, $|A_h \in A : A_i \cap A_h \neq \emptyset| \geq \mu$. They have also used transversal design for key pre-distribution [13]. They have shown that for a prime number p and a integer k such that $2 \leq k \leq p$, there exists a transversal design $TD(k, p)$. In that design, p^2 number of nodes can be arranged with k keys in each node in such a way that (i, j) th node will have the keys $(x, xi + j \text{ mod } p) : 0 \leq x \leq k$. for $0 \leq i \leq p-1$ and $0 \leq j \leq p-1$. If two nodes want to find common keys between them they just need to exchange their node identifiers and the shared key algorithm complexity is $O(1)$.

The communication overhead is $O(\log p) = O(\log \sqrt{N})$ where N is the size of the network. They also gave the estimate of probability of sharing a common key between two nodes and it is $p_1 = k(r-1) / (b-1)$ where k is the keys per node, r is the number of nodes a key is in and b is the total number of nodes in the network. The estimate for resiliency for s node capture is $Fail = 1 - (1 - (r - 2) / (b - 2))^s$. A multiple space has also been presented by Lee and Stinson in [14].

Chakrabarti, Maitra and Roy [5], [6] proposed a hybrid key pre-distribution scheme by merging the blocks in combinatorial designs. They considered Lee and Stinson construction and randomly selected some fixed number of blocks and merged them to form key chains. Though their proposed scheme increased the number of keys per node, it improved the resiliency than Lee and Stinson's Scheme [13].

Simonova, Ling and Wang discuss a homogeneous scheme in [22]. According to them, each grid in the network will have a disjoint key pool. Nodes from the same grid will communicate via this. There will another key pool called deployment key pool which will be constructed from neighboring key pools. Nodes from two neighboring grid can communicate via keys of the deployment key pool. Zhou, Ni and Ravishankar was first to propose a key pre-distribution scheme in [24] where sensors are mobile.

Eschenauer and Gilgor [10] proposed a probabilistic key predistribution scheme to bootstrap the initial trust between the sensor nodes. The main idea was to let each sensor node randomly pick a set of keys from a key pool (mobile polynomial pool) before deployment, so that any two sensor nodes had a certain probability of sharing at least one common key. Chan et al. [9] further extended this idea and developed two key predistribution schemes: the q -composite key predistribution scheme and the random pair wise keys scheme. The q -composite key predistribution scheme also used a key pool, but required two sensor nodes to compute a pair wise key from at least q predistributed keys that they shared. The random pair wise keys scheme randomly picked pairs of sensor nodes and assigned each pair a unique random key. Both schemes improved the security over the basic probabilistic key predistribution scheme.

ACKNOWLEDGMENT

This paper could not be written to its fullest without prof. B.S.Sonawane, who is my guide, as well as one who challenged and encouraged me throughout my time spent studying under him. He would have never accepted anything less than my best efforts, and for that, I thank him.

CONCLUSION

I have seen that, between all above schemes most of the probabilistic schemes are scalable in nature while deterministic schemes are not scalable. But advantage of deterministic schemes is, they are simpler in terms of computation and also better in terms of resiliency and connectivity because of its certainty. Schemes using combinatorial structures are good in terms of resiliency. The basic schemes of Blom or Blundo et al have a good trade-off between storage and security. Mainly in the key management number of schemes are found because they have been researched by various reserchers, they all have some advantages as well as some disadvantages. So the scheme which satisfies both requirements and resources only that scheme we should implement. Security should be a big priority in military services than in civilian application of wireless sensor network. Moreover there are lots of opportunities in this area so that constrained resources of wireless sensor network can be effectively utilized.

REFERENCES:

- [1] Rolf Blom. An optimal class of symmetric key generation systems. In EUROCRYPT, pages 335–338, 1984.
- [2] Carlo Blundo, Alfredo De Santis, Amir Herzberg, Shay Kutten, Ugo Vaccaro, and Moti Yung. Perfectly-secure key distribution for dynamic conferences. In CRYPTO, pages 471–486, 1992.
- [3] Seyit A. C, amtepe and B'ulent Yener. Combinatorial design of key distribution mechanisms for wireless sensor networks. IEEE/ACM Trans. Netw., 15(2):346–358, 2007.
- [4] Seyit Ahmet C, amtepe and B'ulent Yener. Combinatorial design of key distribution mechanisms for wireless sensor networks. In ESORICS, pages 293–308, 2004.

- [5] Dibyendu Chakrabarti, Subhamoy Maitra, and Bimal K. Roy. A key pre-distribution scheme for wireless sensor networks: Merging blocks in combinatorial design. In ISC, pages 89–103, 2005.
- [6] Dibyendu Chakrabarti, Subhamoy Maitra, and Bimal K. Roy. A key pre-distribution scheme for wireless sensor networks: merging blocks in combinatorial design. *Int. J. Inf. Sec.*, 5(2):105–114, 2006.
- [7] Haowen Chan and Adrian Perrig. Pike: peer intermediaries for key establishment in sensor networks. In INFOCOM, pages 524–535, 2005.
- [8] Haowen Chan, Adrian Perrig, and Dawn Song. Random key predistribution schemes for sensor networks. In SP '03: Proceedings of the 2003 IEEE Symposium on Security and Privacy, page 197, Washington, DC, USA, 2003. IEEE Computer Society.
- [9] H. Chan, A. Perrig, and D. Song, “Random Key Pre-Distribution Schemes for Sensor Networks,” *Proc. IEEE Symp. Research in Security and Privacy*, 2003.
- [10] Wenliang Du, Jing Deng, Yunghsiang S. Han, and Pramod K. Varshney. A key predistribution scheme for sensor networks using deployment knowledge. *IEEE Trans. Dependable Sec. Comput.*, 3(1):62–77, 2006.
- [11] Wenliang Du, Jing Deng, Yunghsiang S. Han, Pramod K. Varshney, Jonathan Katz, and Aram Khalili. A pair wise key predistribution scheme for wireless sensor networks. *ACM Trans. Inf. Syst. Secur.*, 8(2):228–258, 2005.
- [12] Laurent Eschenauer and Virgil D. Gligor. A key-management scheme for distributed sensor networks. In CCS '02: Proceedings of the 9th ACM conference on Computer and communications security, pages 41–47, New York, NY, USA, 2002. ACM.
- [13] Jooyoung Lee; D.R. Stinson. A combinatorial approach to key predistribution for distributed sensor networks,. *Wireless Communications and Networking Conference*, 2:1200–1205, 13-17 March 2005.
- [14] Jooyoung Lee and Douglas R. Stinson. On the construction of practical key predistribution schemes for distributed sensor networks using combinatorial designs. *ACM Trans. Inf. Syst. Secur.*, 11(2), 2008
- [15] Donggang Liu and Peng Ning. Establishing pairwise keys in distributed sensor networks. In *ACM Conference on Computer and Communications Security*, pages 52–61, 2003.
- [16] Donggang Liu, Peng Ning, and Wenliang Du. Group-based key predistribution in wireless sensor networks. In *Workshop on Wireless Security*, pages 11–20, 2005.
- [17] Donggang Liu, Peng Ning, and Wenliang Du. Group-based key predistribution for wireless sensor networks. *TOSN*, 4(2), 2008.
- [18] Donggang Liu, Peng Ning, and Rongfang Li. Establishing pairwise keys in distributed sensor networks. *ACM Trans. Inf. Syst. Secur.*, 8(1):41–77, 2005.
- [19] Keith M. Martin, Maura B. Paterson, and Douglas R. Stinson. Key predistribution for homogeneous wireless sensor networks with group deployment of nodes, 2008.
- [20] R. Kannan S.S. Iyengar R. Kalidindi and A. Duresi. Sub-grid based key vector assignment: A key pre-distribution scheme for distributed sensor networks. *Journal of Pervasive Computing and Communications*, 2(1):35–43, 2006.
- [21] Mohammed Golam Sadi, Dong Seong Kim, and Jong Sou Park. Gbr: Grid based random key predistribution for wireless sensor network. In *ICPADS (2)*, pages 310–315, 2005.
- [22] Katerina Simonova, Alan C. H. Ling, and Xiaoyang Sean Wang. Location-aware key predistribution scheme for wide area wireless sensor networks. In *SASN*, pages 157–168, 2006.

- [23] Jooyoung Lee; D.R. Stinson. Common intersection designs,. *Journal of Combinatorial Designs*, 14:251–269, 2006.
- [24]] Li Zhou, Jinfeng Ni, and Chinya V. Ravishankar. Supporting secure communication and data collection in mobile sensor networks. In *INFOCOM*, 2006.
- [25]] Sencun Zhu, Shouhuai Xu, Sanjeev Setia, and Sushil Jajodia. Establishing pairwise keys for secure communication in ad hoc networks: A probabilistic approach. In *ICNP*, pages 326–335, 2003.
- [26] Douglas R. Stinson. *Combinatorial Designs: Construction and Analysis*. Springer-Verlag, 2004