A Unified Approach to Combinatorial Key Predistribution Schemes for Sensor Networks

Douglas R. Stinson

David R. Cheriton School of Computer Science
University of Waterloo

3rd Biennial Canadian Discrete and Algorithmic Mathematics Conference (CanaDAM 2011)
Victoria BC
June 1, 2011

This is joint work with Maura Paterson.

Wireless Sensor Networks

- sensor nodes have limited computation and communication capabilities
- a network of 1000 – 10000 sensor nodes is distributed in a random way in a possibly hostile physical environment
- the sensor nodes operate unattended for extended periods of time
- the sensor nodes have no external power supply, so they should consume as little battery power as possible
- usually, the sensor nodes communicate using secret key cryptography
- a set of secret keys is installed in each node, before the sensor nodes are deployed, using a suitable key predistribution scheme (or KPS)
- nodes may be stolen by an adversary (this is called node compromise)

Two Trivial Schemes

1. If every node is given the same secret master key, then memory costs are low. However, this situation is unsuitable because the compromise of a single node would render the network completely insecure.
2. For every pair of nodes, there could be a secret pairwise key given only to these two nodes. This scheme would have optimal resilience to node compromise, but memory costs would be prohibitively expensive for large networks because every node would have to store $n - 1$ keys, where n is the number of nodes in the WSN.

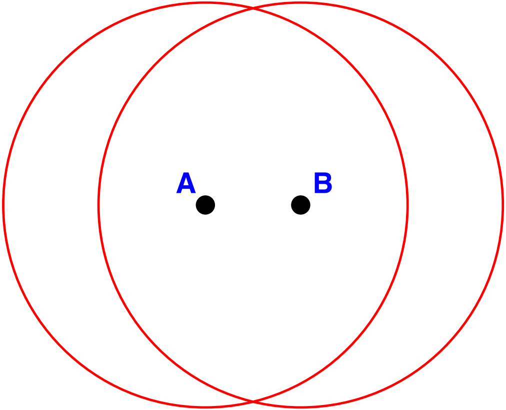
Eschenauer-Gligor and Related Schemes

- In 2002, Eschenauer and Gligor [2] proposed a probabilistic approach to key predistribution for sensor networks. For a suitable value of k, every node is assigned a random k-subset of keys chosen from a given pool of secret keys.
- In 2003, Chan, Perrig and Song [1] suggested that two nodes should compute a pairwise key only if they share at least η common keys, where the integer $\eta \geq 1$ is a pre-specified intersection threshold. Such a pair of nodes is termed a link.
- Suppose that U_i and U_j have exactly $\ell \geq \eta$ common keys, say $\mathbf{key}_{a_1}, \ldots, \mathbf{key}_{a_\ell}$, where $a_1 < a_2 < \cdots < a_\ell$. Then they can each compute the same pairwise secret key,

$$K_{i,j} = h(\mathbf{key}_{a_1} \parallel \ldots \parallel \mathbf{key}_{a_\ell} \parallel i \parallel j),$$

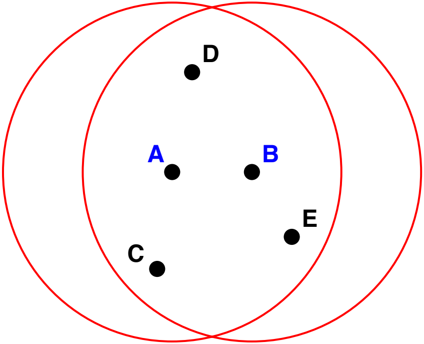
  using a key derivation function h that is constructed from a secure public hash function, e.g., SHA-1.

Multihop Paths



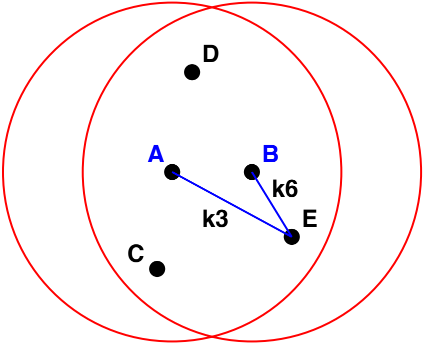
A has keys k1, k3, k5
B has keys k2, k4, k6

Multihop Paths (cont.)



A has keys k1, k3, k5
B has keys k2, k4, k6
C has keys k1, k3, k7
D has keys k2, k6, k7
E has keys k3, k6, k7

Multihop Paths (cont.)



A has keys k1, k3, k5
B has keys k2, k4, k6
C has keys k1, k3, k7
D has keys k2, k6, k7
E has keys k3, k6, k7

Attack Model

- The most studied adversarial model in WSNs is random node compromise [2].
- An adversary compromises a fixed number of randomly chosen nodes in the network and extracts the keys stored in them.
- Any links involving the compromised nodes are broken.
- However, this can also cause other links to be broken that do not directly involve the compromised nodes.
- A link formed by two nodes A_1, A_2, where $|A_1 \cap A_2| \geq \eta$, will be broken if a node $B \notin \{A_1, A_2\}$ is compromised, provided that $A_1 \cap A_2 \subseteq B$.
- If s nodes, say $B_1, \ldots, B_s$, are compromised, then a link A_1, A_2 will be broken whenever

$$A_1 \cap A_2 \subseteq \bigcup_{i=1}^{s} B_i.$$

Important Metrics

Storage requirements

The number of keys stored in each node, which is denoted by k, should be relatively small (e.g., at most 100).

Network connectivity

The probability that a randomly chosen pair of nodes can compute a common key is denoted by $\mathbf{Pr_1}$. $\mathbf{Pr_1}$ should be fairly large (e.g., at least 0.6).

Network resilience

Resilience against node compromise is commonly measured by computing the probability that a random link is broken by the compromise of a single node not in the link. We denote this probability by $\mathbf{fail}$ (high resilience corresponds to a small value for $\mathbf{fail}$).

Remark: As η is increased, $\mathbf{Pr_1}$ and $\mathbf{fail}$ both decrease.

Deterministic Schemes

- In 2004, deterministic KPS were proposed independently by Camtepe and Yener; by Lee and Stinson; and by Wei and Wu.
- A suitable combinatorial design is chosen, and each block is assigned to a node in the WSN (the design and the correspondence of nodes to blocks is public).
- The points in the block are the indices of the keys given to the corresponding node.
- Probabilistic schemes are analyzed using random graph theory, and desirable properties hold with high probability.
- Deterministic schemes can be proven to have desirable properties, and they have more efficient algorithms for shared-key discovery than probabilistic schemes.

Some Proposals for Deterministic Schemes

Projective planes Çamtepe and Yener 2004; Lee and Stinson 2004; Chakrabarti and Seberry 2006.

Generalised quadrangles Çamtepe and Yener 2004.

Configurations Lee and Stinson 2005.

Transversal designs $t = 2$ Lee and Stinson 2005; Chakrabarti and Seberry 2006.

Transversal designs $t = 3$, $\eta = 2$ Lee and Stinson 2005.

Partially balanced incomplete block designs Ruj and Roy 2007.

Spherical geometries Dong, Pei and Wang 2008.

Orthogonal arrays Dong, Pei and Wang 2008; Xu, Chen and Wang 2008.

Reed Solomon codes Ruj and Roy 2008.

Mutually orthogonal latin squares Xu, Chen and Wang 2008.

Rational normal curves in projective spaces Pei, Dong, and Rong 2010.

Comments

- There is considerable duplication of schemes in the above list.
- TDs, OAs, Reed-Solomon codes and MOLS are all essentially the same thing.
- Formulas are developed from scratch in every new proposal for a KPS.
- Perhaps a general, unified approach is warranted.
- Therefore we define a general class of designs that have nice block intersection properties.
- This allows the derivation of general formulas for desired metrics.

Partially Balanced t-designs

- Let v, k, t be positive integers and let λ_i be positive integers, for $0 \leq i \leq t-1$.
- A t-$(v, k, \lambda_0, \ldots, \lambda_{t-1})$-partially balanced t-design (or PBtD) is a pair $(X, \mathcal{A})$ that satisfies the following properties:
  1. $\mathcal{A}$ is a set of k-subsets of X (elements of X are called points and members of $\mathcal{A}$ are called blocks).
  2. There are exactly λ_0 blocks.
  3. For $1 \leq i \leq t-1$, every i-subset of points occurs in either 0 or λ_i blocks.
  4. For $t \leq i \leq k$, every i-subset of points occurs in either 0 or 1 blocks.
- The number of blocks (λ_0) is also denoted by b.
- WLOG, every point occurs in at least one block, so every point occurs in exactly $r = \lambda_1$ blocks.

Examples

- A t-$(v, k, 1)$-design is a t-$(v, k, \lambda_0, \ldots, \lambda_{t-1})$-PB$t$D where
$$\lambda_i = \frac{\binom{v-i}{t-i}}{\binom{k-i}{t-i}}$$
for $0 \leq i \leq t-1$.
- A t-(v, k, λ)-design with $\lambda > 1$ is not necessarily a PBtD. For example, a 2-$(v, 3, 2)$-design is a PBtD if and only if it is a simple design (i.e., a design having no repeated blocks).
- An (s, t)-generalized quadrangle is a 2-$((st+1)(s+1), s+1, \lambda_0, \lambda_1)$-PB$t$D where
$$\lambda_0 = (st+1)(t+1) \text{ and } \lambda_1 = t+1.$$

More Examples

- A TD(t,k,n) is a t-$(kn, k, \lambda_0, \ldots, \lambda_{t-1})$-PB$t$D where
$$\lambda_i = n^{t-i}$$
for $0 \leq i \leq t-1$.
- [4] For a prime power q, the irreducible conics in $\mathrm{PG}(2,q)$ yield a 5-$(q^2+q+1, q+1, \lambda_0, \ldots, \lambda_4)$-PB$t$D where
$$\begin{aligned}\lambda_0 &= q^5 - q^2,\\ \lambda_1 &= q^4 - q^2,\\ \lambda_2 &= q^3 - q^2,\\ \lambda_3 &= q^2 - 2q + 1, \text{ and}\\ \lambda_4 &= q - 2.\end{aligned}$$

Block Intersection Properties of PBtDs

Theorem

Suppose there exists a t-$(v, k, \lambda_0, \ldots, \lambda_{t-1})$-PB$t$D. then for any block B and for any $C \subseteq B$ with $|C| = i \leq t-1$, it holds that

$$|\{A \in \mathcal{A} : A \cap B = C\}| = \mu'(i),$$

where

$$\mu'(t-i) = \sum_{j=0}^{i-1} (-i)^j \binom{k-t+i}{j} (\lambda_{t-i+j} - 1).$$

Remark: For a transversal design (or orthogonal array) with $\lambda = 1$, this is essentially the weight enumerator of the corresponding MDS code.

From PBtD to KPS

- For an integer i such that $\eta \leq i \leq t-1$, an i-link is a set of two blocks $\{A_1, A_2\}$ such that $|A_1 \cap A_2| = i$.
- Let L_i denote the total number of i-links and let

$$L = \sum_{i=\eta}^{t-1} L_i.$$

- Let α_i denote the number of i-links that contain a fixed block A, and let

$$\alpha = \sum_{i=\eta}^{t-1} \alpha_i.$$

- A breaks a link $\{A_1, A_2\}$ if $A \neq A_1, A_2$ and $A_1 \cap A_2 \subseteq A$.
- Let β_i denote the number of i-links that a fixed block A breaks, and let

$$\beta = \sum_{i=\eta}^{t-1} \beta_i.$$

Formulas

Using the λ_i and $\mu'(i)$ values, we can obtain formalas for α_i, β_i and L_i. Then we can compute $\mathbf{fail}$ and $\mathbf{Pr_1}$.

- $\alpha_i = \binom{k}{i}\mu'(i)$.
- $\beta_i = \mu'(i)\left(\frac{\lambda_i}{2} - 1\right)\binom{k}{i}$.
- $L_i = \frac{b\alpha_i}{2}$ and $L = \frac{b\alpha}{2}$.
- $\mathbf{fail} = \frac{\beta}{L - \alpha}$.
- $\mathbf{Pr_1} = \frac{\alpha}{b - 1}$.

Sample Results

scheme	$\mathbf{Pr_1}$	**fail**
$\mathrm{TD}, t = 2$	$\dfrac{k}{n+1}$	$\dfrac{n-2}{n^2-2}$
$\mathrm{TD}, t = 3, \eta = 2$	$\dfrac{k(k-1)}{2(n^2+n+1)}$	$\dfrac{n-2}{n^3-2}$
$\mathrm{TD}, t = 3, \eta = 1$	$\dfrac{k(2n-k+3)}{2(n^2+n+1)}$	$\dfrac{2n^3+(4-2k)n^2+(k-5)n+2k-6}{(2n-k+3)(n^3-2)}$
inv. plane, $\eta = 1$	$\dfrac{n^3+3n^2-2}{2(n^3+n-1)}$	$\dfrac{3n^2+2n-4}{n^4+3n^3+2n^2+2n-4}$
inv. plane, $\eta = 2$	$\dfrac{n^3+n^2}{2(n^3+n-1)}$	$\dfrac{1}{n^2+n+2}$

Asymptotic Results (1)

- We want $\mathbf{Pr_1}$ to be large, but at the same time, we want $\mathbf{fail}$ to be small.
- Given a TD(t, k, n), suppose we fix $k = cn$ and we consider the ratio $\rho = \mathbf{Pr_1}/\mathbf{fail}$ as $n \to \infty$.
- The result has the form dn^j where d is a constant depending on c.
- This provides a convenient single data point to compare different schemes.
- We also consider $k = n + 1$ and $k = n$ as special cases and compare them to schemes based on inversive planes (for $t = 3$) and normal rational curves (for $t = 5$).

Asymptotic Results (2)

scheme	$\mathbf{Pr_1}$	**fail**	ρ
$\mathrm{TD}(3,k,n), \eta = 2, k = cn$	$\frac{c^2}{2}$	$\frac{1}{n^2}$	$\frac{c^2 n^2}{2}$
$\mathrm{TD}(3,k,n), \eta = 2, k = n+1$	$\frac{1}{2}$	$\frac{1}{n^2}$	$\frac{n^2}{2}$
inversive plane, $\eta = 2$	$\frac{1}{2}$	$\frac{1}{n^2}$	$\frac{n^2}{2}$
$\mathrm{TD}(3,k,n), \eta = 1, k = cn, c < 1$	$\frac{c(2-c)}{2}$	$\frac{2(1-c)}{(2-c)n}$	$\frac{c(2-c)^2 n}{4(1-c)}$
$\mathrm{TD}(3,k,n), \eta = 1, k = n$	$\frac{1}{2}$	$\frac{5}{n^2}$	$\frac{n^2}{10}$
$\mathrm{TD}(3,k,n), \eta = 1, k = n+1$	$\frac{1}{2}$	$\frac{3}{n^2}$	$\frac{n^2}{6}$
inversive plane, $\eta = 1$	$\frac{1}{2}$	$\frac{3}{n^2}$	$\frac{n^2}{6}$

Asymptotic Results (3)

Some interesting observations:

1. TD schemes with $t = 3$ and $k = n + 1$ have the same asymptotic behaviour as schemes constructed from inversive planes.
2. Similarly, TD schemes with $t = 5$ and $k = n + 1$ have the same asymptotic behaviour as schemes constructed from normal rational curves (see the next slide).
3. For TD schemes with $\eta = t - 2$, when we set $k = cn$, the value of ρ contains a factor of $(1 - c)$ in the denominator. If we set $k = n$ or $k = n + 1$ in these TDs, then the value of ρ increases by a factor of $O(n)$.
4. In general, ρ increases as η increases.

Asymptotic Results (4)

Asymptotic values of metrics for $\mathrm{TD}(5, n+1, n)$ and NRC-PBtD ($t = 5$):

η	$\mathbf{Pr_1}$	**fail**	ρ
1	$\frac{5}{8}$	$\frac{8}{15n}$	$\frac{75n}{64}$
2	$\frac{7}{24}$	$\frac{6}{7n^2}$	$\frac{49n^2}{144}$
3	$\frac{1}{24}$	$\frac{13}{n^4}$	$\frac{n^4}{312}$
4	$\frac{1}{24}$	$\frac{1}{n^4}$	$\frac{n^4}{24}$

References

[1] H. Chan, A. Perrig, and D. Song. Random key predistribution schemes for sensor networks. In *Proceedings of the 2003 Symposium on Security and Privacy*. IEEE Computer Society, 197–213.

[2] L. Eschenauer and V. Gligor. A key-management scheme for distributed sensor networks. In *Proceedings of the 9th ACM Conference on Computer and Communications Security*. ACM Press, 2002, pp. 41–47.

[3] M. B. Paterson and D. R. Stinson. A unified approach to combinatorial key predistribution schemes for sensor networks. Cryptology ePrint Archive: Report 2011/076, `http://eprint.iacr.org/2011/076`.

[4] D.-Y. Pei, J.-W. Dong and C.M. Rong. A novel key predistribution scheme for wireless distributed sensor networks. *Science China Information Sciences* **53** (2010), 288–298.

thank you for your attention!