Random graphs for WSN security

Osman Yağan and Armand M. Makowski

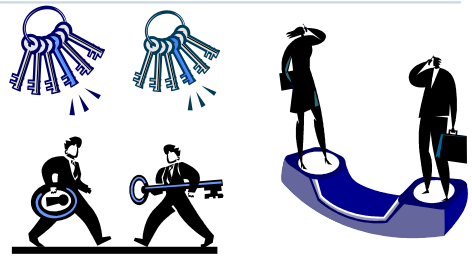
The Institute for Systems Research
A. JAMES CLARK SCHOOL OF ENGINEERING
ISR @25

Wireless sensor networks and security

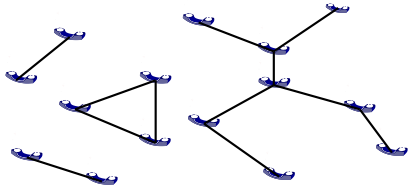
- WSNs are **distributed** collections of sensors with **limited** capabilities for computations and wireless communications.
- Deployed in **hostile** environments where communications are monitored, and nodes are subject to capture and surreptitious use by an **adversary**.
- **Cryptographic protection** needed to ensure secure communications, and to enable sensor-capture detection, key revocation and sensor disabling.
- A proposed solution: Random key predistribution!

A random key predistribution scheme (Eschenauer and Gligor 2002)

- Before deployment, each sensor node is independently assigned K **distinct** cryptographic keys which are selected at **random** from a pool of P keys.
- These K keys form the **key ring** of the node, and are inserted into its memory.
- Any pair of sensors can then establish a secure link between them if they are within transmission range of each other **and** if their key rings have **at least one key in common.**



- However, the resulting network may be **disconnected!**



Q: Given integers P and K with $K \leq P$, how do we select the parameters P and K to make the probability of secure connectivity as large as possible?

The full visibility case: Random key graphs $\mathbb{K}(n;\theta)$

- Full visibility: Sensors are all within communication range of each other.
- n: The number of nodes.
- P: The size of the key pool.
- K: The size of each key ring.
- With $\theta \equiv (P,K)$, let $K_i(\theta)$ denote the **random** set of K **distinct** keys assigned to node i. Assume the random sets $K_1(\theta),\ldots,K_n(\theta)$ to be **i.i.d.** with

$$\mathbb{P}[K_i(\theta) = S] = \binom{P}{K}^{-1}, \quad S \in \mathcal{P}_K$$

- The probability of any pair of sensors having a common key is given by

$$\mathbb{P}\left[K_i(\theta) \cap K_j(\theta) \neq \emptyset\right] = 1 - \frac{\binom{P-K}{K}}{\binom{P}{K}} \simeq \frac{K^2}{P}$$

- Quantity of interest

$$P(n;\theta) \;:=\; \mathbb{P}[\mathbb{K}(n;\theta) \text{ is connected}]$$

Main results

Consider any pair of functions $K, P : \mathbb{N}_0 \to \mathbb{N}_0$ such that $K_n \leq P_n$ for all $n = 1,2,\ldots,n$. Define the sequence $\alpha : \mathbb{N}_0 \to \mathbb{R}$ as the *deviation function* associated with this scaling as

$$\frac{K_n^2}{P_n} = \frac{\log n + \alpha_n}{n}, \quad n = 1,2,\ldots. \tag{1}$$

∗ A zero-one law for connectivity:

Theorem 1 *Consider any admissible pair $K = K_n$ and $P = P_n$ such that $P_n \geq \sigma n$ for some $\sigma > 0$. We have*

$$\lim_{n\to\infty} P(n;\theta_n) = \begin{cases} 0 & \text{if } \lim_{n\to\infty} \alpha_n = -\infty \\ 1 & \text{if } \lim_{n\to\infty} \alpha_n = +\infty \end{cases}$$

∗ The double exponential result:

Theorem 2 *Consider any admissible pair $K = K_n$ and $P = P_n$ such that $P_n \geq \sigma n$ for some $\sigma > 0$. We have*

$$\lim_{n\to\infty} P(n;\theta_n) = e^{-e^{-c}} \tag{2}$$

whenever

$$\lim_{n\to\infty} \alpha_n = c$$

Other application areas of random key graphs

◇ Recommender systems using collaborative filtering - Marbach 2008

- **Netflix**, **Amazon.com**, etc. use recommender systems.
- Users rate the movies they have seen or the products they have purchased.
- A recommender system uses this data to predict the *taste* of each particular user, and to suggest movies or products that they might like.
- Suppose two Netflix subscribers, say Bob and Alice, rate K movies out of the P possible movies provided by Netflix. In recommending a movie for Alice, a recommender system can make use of Bob's movie ratings if they have rated **at least one movie in common.**
- Under the assumption that users rate the same number of movies, random key graphs can be used to model such collaborative-based recommender systems and to evaluate their performance.

◇ Modeling the *small world* effect - Yağan and Makowski 2009

- **Six degrees of separation -** Milgram's experiments suggest that the social network in the United States is **small**.
  - ◇ The path lengths between pairs of individuals are **short**.
- Social networks are highly **clustered -** If Alice and Bob have a common friend, say Carol, it is very likely that Alice and Bob are also friends.
  - ◇ Much more likely than some randomly chosen Alice and Bob being friends.
- Watts and Strogatz: A random graph is considered to be a **small world** if it has **high clustering coefficient** and yet **small average path length.**

∗ Clustering coefficient of random key graphs

Theorem 3 *For any admissible pair $K = K_n$ and $P = P_n$, we have*

$$\mathbb{P}[i \sim j \mid i \sim k \text{ and } j \sim k] \simeq \frac{K^2}{P} + \frac{1}{K}$$

with $i \sim j$ denoting the event that the nodes i and j have at least one key in common and thus linked to each other.

- Recall that $\mathbb{P}[i \sim j] \simeq \frac{K^2}{P}$
  - ◇ K and P can be adjusted so as to satisfy $\frac{K^2}{P} + \frac{1}{K} \gg \frac{K^2}{P}$
  - ◇ Whenever $P \gg K^3$, the random key graph is **highly clustered.**
- Rybarczyk has recently shown that the average path length in a random key graph is **small**.
- Under suitable parameter selection random key graphs are small worlds!