Re-examining Probabilistic Versus Deterministic Key Management

Dahai Xu Jianwei Huang Jeffrey Dwoskin Mung Chiang Ruby Lee
{dahaixu, jianweih, jdwoskin, chiangm, rblee}@princeton.edu
Department of Electrical Engineering, Princeton University, NJ 08544, USA

***Abstract*— It is widely believed that although being more complex, a probabilistic key predistribution scheme is much more resilient against node capture than a deterministic one in lightweight wireless ad hoc networks. Backed up by the surprisingly large successful attack probabilities computed in this paper, we show that the probabilistic approaches have only limited performance advantages over deterministic approaches. We first consider a static network scenario as originally considered in the seminal paper by Eschenauer and Gligor [1], where any node capture happens after the establishment of all pairwise links, and show that the deterministic approach can achieve a performance as good as the probabilistic one. Furthermore in a mobile network, the probabilistic key management as described in [1] can lead to a successful attack probability of one order of magnitude larger than the one in a static network.**

I. Introduction

A. Motivation

Lightweight ad hoc networks typically consist of nodes that are highly distributed with very limited computation and energy resources, such as portable mobile devices and tiny low-cost sensors used for environment surveillance and emergency rescues. As the cornerstone of security communication, various key management schemes have been proposed trying to provide a highly secure communication environment in lightweight ad hoc networks against malicious attacks of adversaries. Among various key management schemes, *symmetric* key predistribution schemes (e.g., [1], [2]) are more suitable to the light weight ad hoc network than *asymmetric* public key schemes, because their resource (e.g., battery, memory, computation power) requirements are small and there is no need for a trusted third party.

There are two main approaches within the symmetric key predistribution schemes: *probabilistic* (e.g., [1]–[5]) and *deterministic* (e.g., [6], [7]). In a probabilistic approach, the keys in each node's key ring are randomly chosen from a large key pool. In a deterministic approach, on the other hand, the key ring is chosen deterministically. In general, probabilistic approaches end up with a large key pool, a larger key ring per node, and poorer network connectivity than the deterministic approaches. [1] On the other hand, a typical deterministic algorithm preloads each node with a single common key and reaches connectivity of 100%. More related references can be found in [8].

It is often believed that a typical probabilistic scheme is much more resilient against node capture than a typical deterministic approach [1], [3], [4], thus making probabilistic schemes popular despite its clear disadvantage on many other metrics when compared with deterministic approach. *In this paper, we show the probabilistic approaches have only limited performance advantages over deterministic approaches*. Our performance measurement is the *Successful Attack Probability* (SAP). An attack on a pairwise link between two authorized nodes is successful if a compromised node can intercept and decipher the information transmitted through that link.

B. Summary of our study between representative probabilistic and deterministic schemes

The probabilistic scheme was first proposed in the seminal and widely cited paper by Eschenauer and Gligor [1], and we call the corresponding scheme the *EG scheme*. It consists of three phases: *key distribution*, *shared key discovery* and *path-key establishment*. In the key distribution phase, each node is loaded with k keys randomly chosen from a large key pool of size m, where $k \ll m$. The shared key discovery is the procedure of establishing a pairwise link between two neighbor nodes if they share one or more key(s). Finally, in the path-key establishment phase, a pairwise link is established between any two neighbor nodes who do not share any key but can establish a path between them through one or more relay nodes. In this case, a path-key is sent from one node to its neighbor through the relays(s), and then a link is established similarly to the shared-key discovery phase.

A representative deterministic scheme is to use single common key, where each node is preloaded with the same initial key. After the deployment, each pair of neighbor nodes exchange messages encrypted by the common initial key to derive a unique (or even random) key for all later communications between them.

Throughout the paper, we will compare the performance of probabilistic and deterministic key management schemes based on the EG scheme [1] and single common key scheme.

We will consider two network scenarios: *static network* and *mobile network*. In a static network, all pairwise links have been established before an adversary can capture any node. This implies that all nodes are deployed almost at the same time and remain static after deployment. This is the case previously considered in [1]. In contrast, in a mobile network, an adversary can capture a node before all pairwise links have been established. This is the case for the network where nodes are constantly on the move and need to establish new links. This includes, for example, a sensor network of buoys floating freely on the ocean to gather environmental data, and a network consisting of sensors moving around in an unknown environment to form a reasonable coverage.

In a static network, the initial common key can be deleted permanently from all nodes after the establishment of all pairwise keys (as in [6]). Therefore, single common key scheme can achieve almost perfect resiliency against node capture (i.e. SAP ≈ 0), since all pairwise keys are randomly generated and known only to the corresponding two neighbor nodes. Thus they cannot be deduced by a captured node even if the common initial key is disclosed. On the other hand, the SAP equals k/m with one captured node in the EG scheme

This work was supported by the National Science Foundation (NSF) under award CNS-0430487.

[1] For example, the probabilistic scheme in [1] requires preloading each node with 83 keys out of a key pool size of $10,000$, and achieves a local direct connectivity of 50%.

1-4244-1429-6/07/$25.00 ©2007 IEEE

where each neighbor node pair uses one of the shared keys to encrypt the communication. The SAP could be reduced to almost zero as in the single common key case if two neighbor nodes also generate a random key for future communication. In short, the deterministic scheme can achieve a performance as good as a probabilistic approach in a static network, but with much lower complexity.

In a mobile network, single common key scheme could lead to an SAP as high as 100% if the common initial key is obtained by an adversary before any link is established. However, we show that the EG algorithm is also quite vulnerable in this case, and may lead to a value of SAP one order of magnitude larger than the one in the static network case (e.g., as high as 60%), especially when the adversary can fully utilize the keys obtained from several compromised nodes. The intuition for the surprising result in this case is as follows. In the static network, there is only one way to attack a link successfully, i.e., knowing the key with which the communications on that link is encrypted. In a mobile network, however, a compromise node can also attack a link by acting as a relay during the path-key establishment phase. By intercepting the key information that is being relayed, a compromised node can figure out the key which the two authorized nodes will use for future mutual communication. This new man-in-the-middle attack opportunity can significantly increase the value of SAP for a probabilistic approach, since there is a high chance of using a relay for link establishment.

The rest of the paper is organized as follows. In Sec. II, we calculate the values of SAP in both static and mobile networks, with a focus on the probabilistic approach (i.e., EG scheme). In Sec. III, we validate the analytical results in Sec. II with simulations based on a C++ simulator and a unit disk network model. We conclude in Sec. IV.

II. Fragility Analysis for Probabilistic Key Management

In this section, we first review the results in [1], where the successful attack probability (SAP) is calculated for a static network. We then consider a mobile network, and show how the value of SAP needs to be substantially revised. We only consider the attacks on the *pairwise* link between two authorized nodes that are within each other's communication range. The SAP will be even higher if A and B are far away and can only be connected with a multi-hop path, since a successful attack on any hop will jeopardize the confidentiality of the whole communication.

The establishment of a link requires two neighbor nodes, A and B, to be able to encrypt the communication over such a link using a common key. This could be achieved in two ways:

(i) A and B share a key within their preloaded key rings, thus can establish the link directly.

(ii) A and B do not share a key initially, and need to exchange additional information through one or more relay nodes, with whom the pairwise links have already been established. For example, A can randomly choose an unused key from its key-ring and send it to B through the relay node(s). Then A and B can use this key to encrypt the pairwise key between them.

In either case, SAP of the link between A and B is defined as

$$SAP \triangleq P(A \otimes B | A \leftrightarrow B), \tag{1}$$

where $A \otimes B$ denotes the event that the link between A and B is successfully attacked, and $A \leftrightarrow B$ denotes the event that A and B establish a link between them. Since a link can only be attacked if it has been established, we have (2) and (3) below.

$$P(A \otimes B \cap A \leftrightarrow B) = P(A \otimes B) \tag{2}$$

$$SAP = \frac{P(A \otimes B)}{P(A \leftrightarrow B)} \tag{3}$$

TABLE I

Summary of Notation

Notations	Meaning
$A \leftrightarrow B$	A and B establish a pairwise link between them
$A \leftrightarrow \mathbb{C}^h \leftrightarrow B$	A and B communicate through one node in $\mathbb{C}^h$
$A \otimes B$	The link between A and B is successfully attacked
$A \sharp B$	A and B share at least one key
$(A \sharp B) \lhd C$	C has all the keys (≥ 1) shared by A and B
$(A \sharp B) \lhd \mathbb{C}^h$	At least one node of $\mathbb{C}^h$ has all the keys (≥ 1) shared by A and B
$(A, B) \sharp \mathbb{C}^h$	At least one node in $\mathbb{C}^h$ shares at least one key with A and at least one key with B
$(A, B) \sharp \mathbb{C}^h_r$	Exactly r nodes out of $\mathbb{C}^h$, each of which shares at least one key with A and at least one key with B

All the notation used in this section are defined in Table I to enable a cleaner presentation of later derivations. A, B and C denote three generic nodes, and $\mathbb{C}^h$ denotes a set of h nodes. Each node is preloaded with a key-ring of k randomly chosen keys out of a key pool of size m.

A. *SAP for a static network*

If a compromised node wants to attack an established link, it needs to know the key used to encrypt the link. Therefore a compromised node can successfully attack an existing link with probability k/m, as stated in [1].

B. *SAP for a mobile network*

In a mobile network, a compromised node C can attack the link between A and B in three ways:

(i) If A and B share a key initially and establish the link directly, then C needs to know the key chosen by A and B to encrypt the link.

(ii) If A and B do not share a key initially and use C as a relay, then C can get the desired information while relaying the information between A and B. A first communicates with C via encrypted messages protected by shared key K_{ac}. C decrypts this with K_{ac} giving it access to the plaintext message, and encrypts this with K_{cb}, a key it shares with node B, then sends the re-encrypted message to B. This sets C up as a man-in-the-middle eavesdropper between A and B, since C can see the plaintext of all messages going from A to B.

(iii) If A and B do not share a key and do not choose C within the relay path, C can still attack the communication between A and B by either eavesdropping on the links along the relay path or attacking the eventual pairwise link established between A and B, if it has any of the keys used for these links.

Overall, the value of SAP depends on the number of compromised nodes and authorized nodes within both A and B's communication range, as well as how A and B choose the relay nodes. To simplify the analysis, we only consider cases (i) and (ii), and further assume only one node relay in case (ii). In the simulation in Sec. III, we calculate SAP for all three cases.

It will be useful to know the probability of sharing at least one key between any two nodes in the network. Denote δ_m^k as the probability that any two nodes A and B do *not* share any key, then

$$\delta_m^k \triangleq P\left(\overline{A\sharp B}\right) = \binom{m-k}{k} \Big/ \binom{m}{k}, \tag{4}$$

where $A\sharp B$ denotes A and B share at least one key. The value of δ_m^k can be either accurately calculated as $\prod_{i=0}^{k-1}(m-k-i)/(m-i)$, or approximated using Stirling's approximation for $n!$ as in [1], i.e.,

$$\delta_m^k = \frac{\binom{m-k}{k}}{\binom{m}{k}} \approx \frac{(1-\frac{k}{m})^{2(m-k+0.5)}}{(1-\frac{2k}{m})^{m-2k+0.5}}. \tag{5}$$

Then the probability of A and B sharing at least one key is

$$P(A\sharp B) = 1 - \delta_m^k. \tag{6}$$

For example, if $k = 83, m = 10000$, $P(A\sharp B) \approx 50\%$.

Next we derive the value of SAP based on the number of authorized users and compromised users within both A and B's communication range. We start with the simplest case, where there is only one compromised node available. We then consider the case where there are h compromised nodes. Finally, we consider the case with h compromised nodes and g authorized nodes.

1) Scenario I: only one compromised node C is within both A and B's communication range: Depending on whether A and B share a key initially, they may establish the pairwise link with or without the relay of C. The probability of successfully establishing the link is (7) and we have (8) below.

$$P(A \leftrightarrow B) = P(A\sharp B) + P((A\sharp C \cap B\sharp C) \cap \overline{A\sharp B}), \tag{7}$$

$$P(A \otimes B) \geq P((A\sharp B) \triangleleft C) + P((A\sharp C \cap B\sharp C) \cap \overline{A\sharp B}). \tag{8}$$

Here $((A\sharp B) \triangleleft C)$ means that A and B share at least one key, and all the shared keys between A and B are within the key-ring of node C. Since we ignore the case where C only knows a subset of the shared keys between A and B, where C still has a chance to successfully attack the link between A and B, we have an inequality in (8) instead of an equality.

Let us calculate each term in (7) and (8). We know the value of $P(A\sharp B)$ from (6). Also,

$$\begin{aligned}
&P(A\sharp C \cap B\sharp C|\overline{A\sharp B}) \\
&= 1 - P(\overline{A\sharp C}) - P(\overline{B\sharp C}) + P(\overline{A\sharp C} \cap \overline{B\sharp C}|\overline{A\sharp B}) && (9)\\
&= 1 - 2\delta_m^k + \binom{m-2k}{k} \Big/ \binom{m}{k} && (10)\\
&= 1 - 2\delta_m^k + \binom{m-k}{k} \Big/ \binom{m}{k} \cdot \binom{m-2k}{k} \Big/ \binom{m-k}{k} && (11)\\
&= 1 - 2\delta_m^k + \delta_m^k \cdot \delta_{m-k}^k && (12)
\end{aligned}$$

Define

$$\phi_m^k \triangleq P(A\sharp C \cap B\sharp C|\overline{A\sharp B}), \tag{13}$$

we then have

$$\begin{aligned}
&P(A\sharp C \cap B\sharp C \cap \overline{A\sharp B}) \\
&= P(\overline{A\sharp B}) \cdot P(A\sharp C \cap B\sharp C|\overline{A\sharp B}) \\
&= \delta_m^k \phi_m^k. && (14)
\end{aligned}$$

Thus from (6), (7) and (14)

$$P(A \leftrightarrow B) = 1 - \delta_m^k + \delta_m^k \phi_m^k, \tag{15}$$

Meanwhile,

$$\begin{aligned}
&P((A\sharp B) \triangleleft C) \\
&= \sum_{i=1}^{k} \left(\binom{k}{i} \cdot \left(\frac{\binom{m-k}{k-i}}{\binom{m}{k}} \right) \cdot \left(\frac{\binom{m-i}{k-i}}{\binom{m}{k}} \right) \right) && (16)\\
&\geq \binom{k}{1} \cdot \left(\frac{\binom{m-k}{k-1}}{\binom{m}{k}} \right) \cdot \left(\frac{\binom{m-1}{k-1}}{\binom{m}{k}} \right) && (17)\\
&= k \left(\frac{k}{m-2k+1} \cdot \frac{\binom{m-k}{k}}{\binom{m}{k}} \right) \cdot \left(\frac{\frac{(m-1)!}{(k-1)!(m-k)!}}{\frac{m!}{k!(m-k)!}} \right) && (18)\\
&= \frac{\delta_m^k k^3}{m(m-2k+1)}, && (19)
\end{aligned}$$

whereas in (17), for simplicity we ignore the event that A, B and C share more than one key. Define

$$\gamma_m^k \triangleq P((A\sharp B) \triangleleft C),$$

we then have

$$SAP = \frac{P(A \otimes B)}{P(A \leftrightarrow B)} \geq \frac{\gamma_m^k + \delta_m^k \phi_m^k}{1 - \delta_m^k + \delta_m^k \phi_m^k}. \tag{20}$$

2) Scenario II: h compromised nodes are within both A and B's communication range: We use $\mathbb{C}^h$ to denote the set of h compromised nodes. Since

$$\begin{aligned}
&P((A,B)\sharp\mathbb{C}^h \cap \overline{A\sharp B}) \\
&= P(\overline{A\sharp B}) \cdot P((A,B)\sharp\mathbb{C}^h|\overline{A\sharp B}) && (21)\\
&= P(\overline{A\sharp B}) \cdot (1 - (1 - P(A\sharp C \cap B\sharp C|\overline{A\sharp B}))^h) && (22)\\
&= \delta_m^k \cdot (1 - (1 - \phi_m^k)^h), && (23)
\end{aligned}$$

then using a similar argument as in Scenario I, we have

$$\begin{aligned}
SAP &\geq \frac{P((A\sharp B) \triangleleft \mathbb{C}^h) + P((A,B)\sharp\mathbb{C}^h \cap \overline{A\sharp B})}{P(A\sharp B) + P((A,B)\sharp\mathbb{C}^h \cap \overline{A\sharp B})} && (24)\\
&\geq \frac{1 - (1 - \gamma_m^k)^h + \delta_m^k \cdot (1 - (1 - \phi_m^k)^h)}{1 - \delta_m^k + \delta_m^k \cdot (1 - (1 - \phi_m^k)^h).} && (25)
\end{aligned}$$

3) Scenario III: h compromised nodes and g authorized nodes are within both A and B's communication range: In this case, if A and B do not share any key initially and need to communicate through a relay, a successful attack can happen if one compromised node is chosen as the relay. Assuming there are a total of *a qualified* relays (i.e, nodes who can establish pairwise links with both A and B), b out of which are compromised nodes. Denote μ_a^b as the probability of A and B picking a compromised node as the relay, which could be b/a under honest attacks, or some higher value under smart attacks.

The probability of having r useable relays out of all h compromised nodes when A and B do not share keys is

$$P((A,B)\sharp\mathbb{C}^h_r|\overline{A\sharp B})$$

$$= \binom{h}{r}\left(P(A\sharp C\cap B\sharp C|\overline{A\sharp B})\right)^r\left(1-P(A\sharp C\cap B\sharp C|\overline{A\sharp B})\right)^{h-r} \tag{26}$$

$$= \binom{h}{r}(\phi^k_m)^r(1-\phi^k_m)^{h-r}. \tag{27}$$

Similarly, the probability of having w useable relays out of all g authorized nodes when A and B do not share keys is

$$P((A,B)\sharp\mathbb{C}^g_w|\overline{A\sharp B}) = \binom{g}{w}(\phi^k_m)^w(1-\phi^k_m)^{g-w}. \tag{28}$$

Then the probability of sending a message through a compromised node given the existence of h compromised nodes, g authorized nodes and A, and B do not share any key is

$$P(A\leftrightarrow\mathbb{C}^h\leftrightarrow B|\overline{A\sharp B})$$

$$= \sum_{r=1}^{h}\sum_{w=0}^{g}\mu^r_{r+w}(P((A,B)\sharp\mathbb{C}^h_r|\overline{A\sharp B})\cdot P((A,B)\sharp\mathbb{C}^g_w|\overline{A\sharp B})) \tag{29}$$

$$= \sum_{r=1}^{h}\sum_{w=0}^{g}\mu^r_{r+w}\left(\binom{h}{r}\binom{g}{w}\left((\phi^k_m)^{r+w}(1-\phi^k_m)^{h+g-(r+w)}\right)\right). \tag{30}$$

Since

$$P(A\leftrightarrow\mathbb{C}^h\leftrightarrow B\cap\overline{A\sharp B}) = P(A\sharp B)\cdot P(A\leftrightarrow\mathbb{C}^h\leftrightarrow B|\overline{A\sharp B}), \tag{31}$$

we have the following lower bound on SAP

$$SAP = \frac{P(A\otimes B)}{P(A\leftrightarrow B)} \tag{32}$$

$$\geq \frac{P((A\sharp B)\lhd\mathbb{C}^h)+P(A\leftrightarrow\mathbb{C}^h\leftrightarrow B\cap\overline{A\sharp B})}{P(A\sharp B)+P(A\leftrightarrow\mathbb{C}^{h+g}\leftrightarrow B\cap\overline{A\sharp B})} \tag{33}$$

$$= \frac{1-(1-\gamma^k_m)^h+\delta^k_m\cdot\left(\sum_{r=1}^h\sum_{w=0}^g\mu^r_{r+w}\left(\binom{h}{r}\binom{g}{w}\left((\phi^k_m)^{r+w}(1-\phi^k_m)^{h+g-(r+w)}\right)\right)\right)}{1-\delta^k_m+\delta^k_m\cdot\left(1-(1-\phi^k_m)^{h+g}\right)}. \tag{34}$$

4) Numerical results: Table II shows the SAP for different values of h and g based on the previous analysis. The key-ring size is $k = 83$, with a key pool size of $m = 10000$.

TABLE II

Successful attack probability (SAP) for different numbers of authorized nodes (g) and compromised nodes (h). We assume there are a total of a qualified relays, b out of which are compromised nodes. μ^b_a is the probability of picking a compromised node as the relay. The key pool size $m = 10000$, the preloaded key-ring size $k = 83$, and the original SAP estimation is hk/m.

h	$g = 0$	$g = 10$		$g = 20$		hk/m
		$\mu^b_a = b/a$	$\mu^b_a = 1$	$\mu^b_a = b/a$	$\mu^b_a = 1$	
1	20.4%	4.7%	13.0%	2.7%	12.8%	0.8%
2	31.1%	8.8%	22.7%	5.1%	22.4%	1.7%
3	37.6%	12.3%	30.0%	7.4%	29.7%	2.5%
4	41.9%	15.3%	35.5%	9.5%	35.2%	3.3%
5	44.8%	18.0%	39.7%	11.4%	39.5%	4.2%
6	46.9%	20.4%	42.9%	13.2%	42.7%	5.0%
7	48.5%	22.5%	45.4%	15.0%	45.2%	5.8%
8	49.7%	24.4%	47.3%	16.6%	47.2%	6.6%
9	50.6%	26.2%	48.8%	18.1%	48.7%	7.5%

Several observations are in order. When (the probability of picking a compromised node as the relay), $\mu^b_a = b/a$, the SAP increases with h (compromised nodes) under fixed g (authorized nodes). When $\mu^b_a = 1$, the general trend is similar, but the SAP is not very sensitive in the cases of $g = 10$ and $g = 20$, since A and B will always choose a compromised node as relay if possible. Comparing with the value of SAP estimated in [1], which is approximated as hk/m, the SAP in Table II is much larger. For example, with $\mu^b_a = b/a$, $h = 9$ and $g = 20$, we have a SAP of 18.1%, as opposed to $hk/m = 7.5\%$. The value of SAP increases further when $\mu^b_a = 1$.

The value of μ^b_a heavily depends on the attack model used by the compromised nodes. Two attack models, *honest attack* and *smart attack*, are defined in Sec. I. In an honest attack, the relays nodes are randomly chosen and $\mu^b_a = b/a$. In a smart attack, however, the compromised nodes will improve the value of μ^b_a by various methods. In a *smart attack with incentive*, the compromised nodes provide incentives for nodes A and B to choose one of them as relay. If the choice of relay is determined by a shortest path routing protocol, the compromised nodes can announce distance metrics of the links connected to them smaller than the actual values. If the choice of relay is based on energy efficiency, the compromised nodes can pretend to be very energy efficient. In most cases, the incentives provided by the compromised nodes can make the value of μ^b_a very close to 1. In *a smart attack with virtual node fabrication*, each compromised node is able to collect the keys from all other compromised nodes, then can fabricate up to $\binom{hk}{k}$ nodes with distinct key rings. The number will be very large if $h \geq 2$. For example, when two nodes are captured with non-overlapping key rings, then

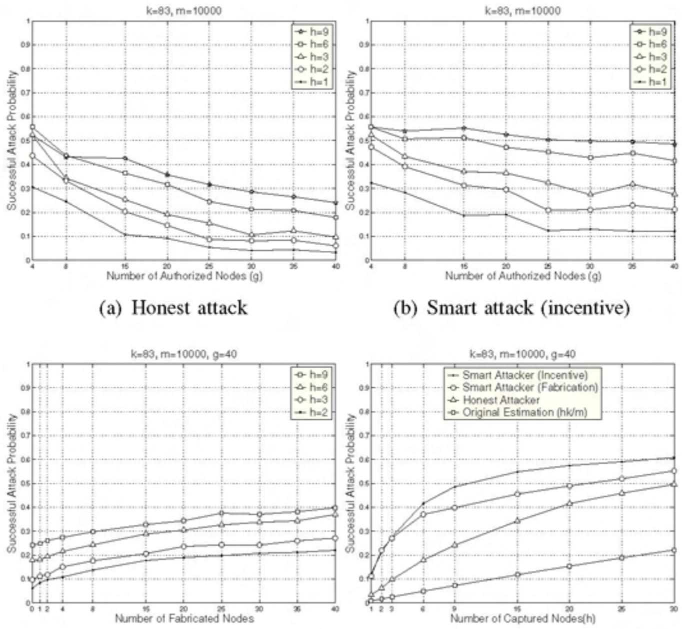
$$\binom{2k}{k} = \frac{(2k)!}{k!} \approx \frac{\sqrt{2\pi}(2k)^{2k+0.5}e^{-2k}}{(\sqrt{2\pi}(k)^{k+0.5}e^{-k})^2} = \frac{2^{2k+0.5}}{\sqrt{2\pi k}}, \tag{35}$$

which is around 5.8×10^{48} if $k = 83$. As a result, the value of μ^b_a will be closer to 1 with the increase of the number of fabricated nodes.

III. Simulation Results

To verify our probability computations in Sec. II, we evaluate the SAP of the probabilistic key predistribution scheme (the EG scheme) through a simulator written in C++. We consider a unit disk network model. A total of g authorized nodes are uniformly distributed in the unit disk. All the compromised nodes (including any virtually fabricated nodes) are placed at the center of the unit disk. All nodes are assumed to have the same transmission range equal to the radius of the disk. This means an adversary can eavesdrop on any communication in the unit disk through the compromised nodes as long as it has the right key(s). Two neighbor nodes will setup a pairwise link if they share one or more keys. Otherwise, they will try to find an relay path through one or more nodes to exchange additional key information, so that they can set up pairwise link between them. When there is more than one qualified relay node available, the authorized nodes will choose a relay randomly in the case of $\mu^b_a = b/a$ (i.e, honest attack or finite virtual node fabrication), or search for a shortest relay path in an attack with incentive. [2] Any two nodes that are not neighbors cannot establish pairwise links among themselves. The main reason of using the above unit disk network model

[2] In the simulation, the smart attack with incentive is approximated as setting the cost of the links adjacent to the compromised nodes as 0.9999 instead of as 1 unit (hop) for other authorized nodes.



(a) Honest attack

(b) Smart attack (incentive)

(c) Smart attack (node fabrication)

(d) Different attack models

Fig. 1. Successful Attack Probability with various numbers of captured nodes (h) and authorized nodes (g)

is to derive a uniform and fair metric (SAP) among various approaches where failing to attack is only due to the lacking in appropriate keys rather than the limitation of transmission range.

The SAP is calculated as the fraction of the links that can be eavesdropped by the compromised nodes among all the pairwise links. As we explained in Sec. I, a basic deterministic scheme like single common key either enables almost zero SAP in a static network, or leads to 100% SAP for the unit disk model in a mobile network since it has the common key and can observe all the key exchanges between nodes. Hence, our focus here is to determine the SAP for the probabilistic key predistribution scheme (i.e., the EG scheme). All the simulation results are averaged over 10 sets of random seeds which affect the distribution of the authorized nodes within the unit disk, the key ring preloaded to each node and the choices in case of multiple qualified relays.

Figs. 1(a) to 1(d) illustrate the values of SAP under different assumptions on the number of compromised nodes (h), number of authorized nodes (g) and different attack models (honest attack, smart attack with incentive, or smart attack with fabrication).

Fig. 1(a) shows the SAP for various values of h and g under the *honest attack*. For a fixed value of h, the SAP decreases when the density of authorized nodes increases. This is because in a denser network, there are more qualified relay nodes available between any two neighbor nodes, thus the probability of choosing a compromised node as the relay is smaller under honest attack. For a fixed number of authorized nodes g, a higher value of h increases the probability of picking a compromised node as the relay, thus leads to a higher value of SAP. In a network where there are 9 compromised nodes and 15 authorized nodes, the SAP could be as high as 42%.

Fig. 1(b) shows the SAP for various values of h and g under the *smart attack with incentive*. In this case, two neighbor nodes without a common key will have high chance to pick a compromised node as relay if it is qualified. There is a high probability of finding a qualified relay node among the compromised nodes when h is large, in which case the SAP is insensitive to the number of authorized nodes g. In a network with 40 authorized nodes and 9 compromised nodes, the SAP would be around 50%.

Fig. 1(c) shows the SAP for the smart attack of various numbers of compromised nodes and different total numbers of virtually fabricated nodes. The total number of authorized nodes is kept at 40. The node fabrication is achieved as follows. All the keys collected from the h compromised nodes will constitute a *compromised key pool*. Then each fabricated node will be loaded with $k = 83$ keys randomly chosen from the compromised key pool. A larger number of fabricated nodes increases the chance of such a node being chosen as a relay node, thus increasing SAP. A larger value of h leads to a larger compromised key pool, which again increases the chance of a fabricated node serving as a qualified relay.

Fig. 1(d) shows the SAP under different numbers of captured nodes, for different kinds of attacks as well as the estimation based on [1] [3]. The number of authorized nodes is fixed at 40. It is clear that the results in [1] significantly underestimate the SAP in mobile networks. With a large enough number of compromised nodes, the SAP can easily reach an unacceptably high value of 50% with all attack models.

IV. Concluding Remarks

In this paper, we discuss the key management in lightweight mobile ad hoc networks. Backed up by the large successful attack probabilities computed in this paper, we show that the probabilistic key predistribution schemes are in fact quite vulnerable to node captures in many practical cases. Considering the large key pool and key ring sizes, complex key predistribution, low network connectivity, and complex pairwise link establishments, the advantage of the probabilistic approach over the deterministic approach is not as much as people have believed.

References

[1] L. Eschenauer and V. D. Gligor, "A key-management scheme for distributed sensor networks," in *CCS'02*, New York, NY, 2002, pp. 41–47.

[2] R. D. Pietro, L. V. Mancini, and A. Mei, "Random key-assignment for secure wireless sensor networks," in *SASN'03*, New York, NY, 2003, pp. 62–71.

[3] H. Chan, A. Perrig, and D. Song, "Random key predistribution schemes for sensor networks," in *IEEE Symposium on Security and Privacy*, 2003.

[4] W. Du et al., "A key management scheme for wireless sensor networks using deployment knowledge," in *INFOCOM'04, Hong Kong*, Mar. 2004.

[5] J. Hwang and Y. Kim, "Revisiting random key pre-distribution schemes for wireless sensor networks," *ACM workshop on Security of ad hoc and sensor networks*, pp. 43–52, 2004.

[6] S. Zhu, S. Setia, and S. Jajodia, "LEAP: efficient security mechanisms for large-scale distributed sensor networks," in *CCS'03*, New York, NY, 2003, pp. 62–72.

[7] J. Lee and D. Stinson, "Deterministic key predistribution schemes for distributed sensor networks," *Selected Areas in Cryptography*, 2004.

[8] S. A. Çamtepe and B. Yener, "Key distribution mechanisms for wireless sensor networks: a survey," Rensselaer Polytechnic Institute, Computer Science Department, Tech. Rep. TR-05-07, Mar. 2005, available at http://www.cs.rpi.edu/research/pdf/05-07.pdf.

[3] When the network is static, an adversary captures h nodes, then its successful attack probability on a link is $1-(1-\frac{k}{m})^h \approx \frac{hk}{m}$.