Distributed and Secure Bootstrapping of Mobile Ad Hoc Networks: Framework and Constructions

SHOUHUAI XU
University of Texas
and
SRDJAN ČAPKUN
ETH Zurich

Secure bootstrapping of mobile ad hoc networks (MANETs) is a challenging problem in scenarios in which network users (or nodes) do not share trust relationships prior to the network deployment. In recent years, a number of schemes have been proposed to solve this problem, assuming either no or limited trust between the nodes prior to their deployment. Despite numerous proposals, there is no common understanding of the proposed schemes and of the trade-offs that they provide. This has consequences for both researchers and practitioners, who do not have a clear idea how to compare the schemes and how to select a scheme for a given application. In this article, we present a framework that helps in understanding and comparing schemes for secure bootstrapping of MANETs. The framework is general because it is policy-neutral and can accommodate many existing bootstrapping schemes. The proposed framework can equally serve as a good basis for the development of new MANET bootstrapping schemes; we show how the development of the framework leads to two new (classes of) distributed bootstrapping schemes. Within the framework, we not only investigate and characterize the properties of the relevant bootstrapping schemes, but also give methods for practitioners to select the relevant system parameters in the Random Walk and the (Restricted) Random Waypoint mobility models.

Categories and Subject Descriptors: C.2.4 [**Computer-Communication Networks**]: Distributed Systems

General Terms: Security

Additional Key Words and Phrases: MANETs, security bootstrapping, secure communication

The work of Shouhuai Xu was supported in part by ARO and UTSA CIAS. The views and conclusions contained in the article are those of the authors and should not be interpreted as, in any sense, the official policies or endorsements of the government or the agencies.
Authors' addresses: S. Xu, Department of Computer Science, University of Texas, San Antonio, San Antonio, TX 78249; email: shxu@cs.utsa.edu; S. Čapkun, Department of Computer Science, ETH Zurich 8092 Zurich, Switzerland; email: capkuns@inf.ethz.ch.

ACM Reference Format:
Xu, S. and Čapkun, S. 2008. Distributed and secure bootstrapping of mobile ad hoc networks: Framework and constructions. Trans. Inf. Syst. Secur. 12, 1, Article 2 (October 2008), 37 pages. DOI = 10.1145/1410234.1410236. http://doi.acm.org/10.1145/1410234.1410236.

1. INTRODUCTION

1.1 Motivation

Bootstrapping Mobile Ad Hoc Networks (MANETs) is an important and challenging problem that has been investigated by many researchers so far. Given the diversity of future MANET applications and of their deployment scenarios, assumptions that are made in the development of their bootstrapping schemes vary. We observe that most existing MANET bootstrapping schemes make some of the following *strong* assumptions:

—The network users are known and present when the network is formed so that an authority can admit them into the network once and for all. Many schemes [Eschenauer and Gligor 2002; Chan et al. 2003; Zhu et al. 2003; Du et al. 2003; Liu and Ning 2003; Camtepe and Yener 2004; Lee and Stinson 2004; Chan and Perrig 2005] fall into this category. This assumption may be appropriate for sensor networks, which may be deployed and controlled by a single authority. However, it may not be suitable for MANETs, where we do not necessarily know which users will become network nodes until the network is formed.

—The users already hold public key certificates of (a subset of) other users before a MANET is formed; schemes presented in Capkun, Buttyan, and Hubaux [2003] and Capkun, Hubaux, and Buttyan [2003] fall into this category. While this may be true for some application scenarios, there are many scenarios in which users may not possess other users' public key certificates or other similar credentials. In these scenarios, it would be useful for users to be admitted into a network just based on their photo identities such as driver's licences. This flexibility may be seen as an advantage of MANETs.

—A MANET can provide secure routing even before security associations have been established; schemes based on threshold public key cryptosystems [Zhou and Haas 1999; Yi and Kravets 2003] fall into this category. These schemes suffer from the "routing-security interdependence cycle" problem, which is essentially the chicken-and-egg problem concerning whether *secure routing* or *secure associations* should be established first. We note that secure routing is possible only after security associations have been established [Hu et al. 2002; Papadimitratos and Haas 2002, 2003]. We also note that, in some special scenarios, the chicken-and-egg problem may be absent (e.g., any joining node can always find a threshold number of nodes within one-hop distance), but another perhaps more serious problem we call "proximity-caused insecurity" emerges (see next item).

—There are a few schemes that do not necessarily suffer from the above "routing-security interdependence cycle" problem. However, such schemes,

including the aforementioned special cases of those based on threshold cryptosystems [Zhou and Haas 1999; Kong et al. 2001; Yi and Kravets 2003; Xu and Iftode 2004], might not offer the required security guarantees, due to the following realistic threat of "proximity-caused insecurity": In order to avoid the routing-security interdependence cycle problem, there are always a threshold number of authorized users that are physically close to each other (i.e., within one-hop communication distance so that routing is eased). As a consequence, it is possible that an adversary compromises these nodes within a short period of time (e.g., by capturing the nodes and/or compromising them one by one).[1]

The above discussion suggests that the MANET bootstrapping problem is not yet entirely solved and that there is no common understanding of this problem. We are not aware of any scheme that can get rid of all the strong assumptions mentioned above. This might have been caused by the lack of a framework using which researchers can deepen their understanding of, and compare different, MANET bootstrapping schemes.

1.2 Our Contributions

We make two contributions in this work. First, we propose a framework that helps in better understanding of the problem of distributed and secure bootstrapping of MANETs. The framework does not assume that the users admitted to the MANET are known in advance. Moreover, users can be admitted into a MANET in a distributed fashion, meaning that no (real-time) interactions between the authorities (called initiators) are necessary during the process of admitting users. The framework is general enough to accommodate existing schemes. This allows us to compare these schemes with respect to a common base. Furthermore, the framework is policy-neutral because it distills mechanisms from policies, under which users are admitted (e.g., how many "approvals" are needed before an outside user is admitted into a MANET; which cryptographic credentials or which picture ID cards does a user need to present in order to get admitted). This means that our framework can accommodate a large class of policies.

[1]It is worthwhile to note that this problem might not be solved by assuming that all the users have tamper-resistant hardware modules due to the following reasons. First, the users to be admitted are not known in advance and thus may not have any such hardware modules. Second, even if they do, the users themselves are not necessarily trusted (e.g., they could use some corrupt hardware modules), which are certainly breakable even by attackers with reasonable hardware skills [Anderson and Kuhn 1996].

It is also worthwhile to mention that proactive cryptosystems [Ostrovsky and Yung 1991], which were introduced in the context of distributed computing in wired networks, might not be able to resolve this problem because a threshold number of users, who are geographically close to each other, could be captured within a short period of time. Moreover, proactive cryptosystems incur heavy computations and communications, which put their practicality in the setting of MANETs in question [Narasimha et al. 2003]. This problem served as a motivation for the following work [Saxena et al. 2005].

Second, since the framework can help us identify new bootstrapping schemes, we propose two new (classes of) practical schemes. We analyze and characterize the properties of the schemes, and investigate methods for practitioners to select system parameters in the Random Walk and (Restricted) Random Waypoint mobility models. We stress that, unlike existing schemes, the new schemes avoid both the routing-security interdependence cycle problem and the proximity-caused insecurity problem. The former problem is prevented by exploiting mobility [Capkun et al. 2003] to allow a pairing method [Hoepman 2004; Čagalj et al. 2006] for admitting users and issuing them cryptographic credentials. The latter problem is prevented because, in order to compromise a whole MANET, the adversary has to compromise all the initiators (or the large number of dispersed users). Fortunately, the initiators are not within the one-hop proximity of each other, unless the underlying mobility model imposes this (which would occur with a small probability in a random mobility model). Furthermore, two techniques are used to help address the small but nonzero probability that the initiators are within the one-hop distance of each other.

(1) The initiators are equipped with some tamper-resistant hardware modules. Note that this idea is applicable here because the (small number of) initiators are trusted. As a result, capturing an initiator may not expose the relevant cryptographic keys to the adversary.
(2) Given that tamper-resistance is still a heuristic security notion [Anderson and Kuhn 1996], it must be used with caution. For this, we allow the initiators to automatically downgrade their roles to normal users (e.g., by securely erasing all of the secrets beyond necessary to act as normal users) in a distributed fashion (i.e., without requiring any coordination). This further reduces the chance that the adversary captures all the initiators, because the cryptographic keys corresponding to their initiator roles are erased after a short period of time.

The rest of the article is organized as follows. In Section 2, we explore the framework. In Section 3, we present an instantiation of the framework based on public key cryptosystems. In Section 4, we present a class of instantiations based on symmetric key systems. In Section 5, we show how a practitioner should select parameters in practice. In Section 6 we discuss related works. We conclude the article in Section 7 with some challenging open problems.

2. DISTRIBUTED AND SECURE BOOTSTRAPPING FRAMEWORK FOR MANETS

In this section we first specify the system model and the communication channel model upon which the framework is built. Then we explore the framework through its components and properties. Finally we show how the framework accommodates existing secure bootstrapping schemes, and discuss its usefulness.

The main notations used throughout this article are summarized below.

T, t	discrete system time $0, 1, \ldots$
$\mathbb{I}_T$	the set of initiators at time T, where $\mathbb{I}_T = \{I_{1_T}, \ldots, I_{\ell_T}\}$
$\mathbb{I}$	the set of all initiators, namely $\mathbb{I} = \mathbb{I}_1 \cup \ldots \cup \mathbb{I}_T$ (when T is no factor, simply $\mathbb{I} = \{I_1, \ldots, I_\ell\}$)
s_T	the threshold number of initiators needed to admit a user at time T where $1 \leq s_T \leq \ell_T$ (in the case T is no factor, simply $1 \leq s \leq \ell$)
$\mathbb{U}_T$	the set of users admitted at time T, where $\mathbb{U}_T = \{U_{1_T}, \ldots, U_{n_T}\}$
$\mathbb{U}$	the set of all admitted users, namely $\mathbb{U} = \mathbb{U}_1 \cup \ldots \cup \mathbb{U}_T$ (when T is no factor, simply $\mathbb{U} = \{U_1, \ldots, U_n\}$)
Γ	system parameters of cryptosystems of the initiators
$(\mathsf{PK}_i, \mathsf{SK}_i)$	the pair of public and private keys of initiator $I_i \in \mathbb{I}$
(pk_U, sk_U)	the pair of public and private keys of admitted user $U \in \mathbb{U}$
$f_{ij}(x, y)$	a bivariate polynomial common to initiators I_i and I_j
$t_M^{(n\times s;\ell)}$	the time at which each of the n users has obtained s credentials from s (out of the ℓ) initiators
$\overline{t_M}^{(n\times s;\ell)}$	mean of $t_M^{(n\times s;\ell)}$

2.1 System Model

A MANET may be formed on-the-fly to accomplish some task(s). Therefore, it would be reasonable to assume that some *initiators* would bootstrap a MANET by admitting outside users according to a predetermined policy, denoted by policy, and by issuing some cryptographic credentials to the admitted users. For example, a policy may specify: (1) the criteria for admitting users such as the possession of a valid driver's license; (2) whether admitted users are allowed to admit other outside users; (3) s_T, the number of approvals an outsider user needs to obtain from the initiators belonging to $\mathbb{I}_T$ at time T (therefore, $1 \leq s_T \leq \ell_T$). Since the specification of a policy depends on the application semantics, we treat it as a black-box procedure without discussing it any further in this article. It is important to realize that the outside users to-be-admitted are not known in advance; instead, any outside user can be admitted as long as it satisfies the policy and there is still a need to admit new users.[2]

In general, we consider a discrete time mode with finite system time $T = 0, 1, \ldots$. Let $\mathbb{I}_T = \{I_{1_T}, \ldots, I_{\ell_T}\}$ be the initiators at time T, and $\mathbb{U}_T = \{U_{1_T}, \ldots, U_{n_T}\}$ be the identities assigned to the users admitted at time T.

[2]However, we are conservative about allowing admitted users to admit other users—a policy that has been advocated in existing threshold cryptosystem based bootstrapping schemes. Regardless of the trust issue (i.e., whether we should put such a strong trust on the admitted users), there are significant security consequences. Specifically, in order to counter an adversary that may launch the Sybil attack or may compromise nodes, the number of approvals—or the threshold in threshold cryptosystem based schemes—must be sufficiently large; otherwise, the adversary could relatively easily compromise the whole system. However, a sufficiently large threshold will not only make the resulting schemes much less efficient, but also trap into the "routing-security interdependence cycle" problem (because it is unlikely that a large threshold number of nodes will always be within one-hop distance). Again, as mentioned before, even if there are more than a threshold number of nodes within one-hop distance, the "proximity-caused insecurity" problem arises.

Denote by $\mathbb{I} = \mathbb{I}_0 \cup \mathbb{I}_1 \cup \ldots$ the population of initiators, and by $\mathbb{U} = \mathbb{U}_0 \cup \mathbb{U}_1 \cup \ldots$ the population of admitted users. We assume that $\mathbb{I} \cap \mathbb{U} = \emptyset$. However, it is reasonable that an initiator $I_{i_0} \in \mathbb{I}_0$ is also admitted by itself or by other initiators at $T = 0$. As a result, I_{i_0} possesses two credentials: one corresponds to its role of being an initiator, and the other corresponds to its role of being a normal user. (Exactly the cryptographic keys corresponding to the initiator role will be erased when an initiator downgrades to become a normal user.)

We may assume the existence of some tamper-resistant hardware modules, such as Trusted Platform Modules (TPMs) specified by the Trusted Computing Group [TCG]. The modules are manufactured by different vendors so as to avoid any homogeneous vulnerability (i.e., each initiator possesses a different type of tamper-resistant hardware module). Since the notion of tamper-resistance is still heuristic, and the number of different types of such hardware modules is relatively small, we must utilize them with cautions. In other words, we should not base security of the whole system exclusively on the tamper-resistance of such hardware modules.

2.2 Communication Channel Assumption

To accommodate the worst-case scenario, we assume that the adversary has total control over all the normal wireless channels. However, in order to facilitate the authentication of outside users and the issuance of credentials to the admitted ones, we assume that there are some special channels beyond the normal wireless channels. Specifically, we assume the existence of an *authenticated private channel* between two users when they move into within a short distance of each other. Such a channel simultaneously ensures *source identification* (i.e., no impersonation of a peer user is possible), *channel integrity* (i.e., no tamperation of message is possible), and *channel confidentiality* (i.e., no information of a transmitted message is leaked to any one other than the two peer users). This remains to be true even if the two users are in possession of no existing (direct or indirect) trust in each other. Authenticated private channels can be based on physical contact or infrared ones [Stajano and Anderson 1999; Balfanz et al. 2002]. Recent advancement shows that they can even be based on radio links. While radio links facilitate better usability, they are potentially more vulnerable. Fortunately, the vulnerability can be addressed by letting the users type a password into their respective devices [Asokan and Ginzboorg 2000], compare strings of words (the longer the string, the higher the security) [Hoepman 2004; Čagalj and Hubaux 2004], exploit the physical proximity of the entities [Čagalj et al. 2006] via distance-bounding [Brands and Chaum 1994], or make use of the characteristics of the wireless channel to provide integrity [Čagalj et al. 2006].

In some cases (e.g., in the public key cryptosystem based scheme that will be detailed later), we may only need a weaker communication channel assumption, namely the existence of a authenticated channel between two users when they move into within a short distance of each other. An authenticated channel assures source identification and channel integrity, but not channel confidentiality. Such channels are also possible because authentication is

straightforward as users within a short distance can visually recognize each other (see, e.g., McCune et al. [2005]), or even when they meet for the first time, they can be introduced to each other by a common friend in whom they trust or by checking each other's identity. Some even more recent results include Goodrich et al. [2006].

Note that when authentication private channels are necessary but only authenticated channels are available, we can build authenticated private channels using authenticated channels as follows. Two users can obtain each other's (perhaps one-time) public key over the authenticated channel; the public keys are then used to support a standard authenticated key exchange protocol (e.g., Krawczyk [2003]) to establish common secrets, which are then used to enforce channel confidentiality etc. over the (normal wireless) channels. Moreover, even if the authenticated string is short (e.g., 20 bits), which may be pertinent to the narrow bandwidth of the special channel, secure communication is still possible [Vaudenay 2005].

In summary, the framework assumes the existence of authenticated (private) channels between two users when they are within a short distance of each other. When authenticated private channels exist, the resulting solutions can be more efficient than the ones that are solely based on authenticated channels. We notice that communications may be conducted fully over the afore-discussed authenticated (private) channels, or partially over them because of their potentially narrow bandwidth (i.e., there are only used to establish common cryptographic keys that are then used to protect the communications over the normal wireless channels). We note that this exploitation of mobility for security is practical because physical presence is perhaps the best way to increase mutual trust and to exchange information in a secure way, especially in the context of MANETs (e.g., man-in-the-middle attack becomes infeasible).

2.3 Framework

The core of the framework consists of the following four processes: (1) *setup of initiators*; (2) *system self-forming*, during which outside users are admitted and may conduct secure communications; (3) *optional initiator downgradation*, during which the initiators may automatically downgrade their roles to normal users; and (4) *establishment of security associations*. To highlight the basic idea underlying the framework, in the presentation of the core of the framework we treat the set of initiators as a fix set, although the number of initiators may be dynamic (as discussed in Section 2.1).

(1) **Setup of initiators**. Each initiator $I \in \mathbb{I}$ establishes an appropriate credential scheme (e.g., a public key cryptosystem for issuing certificates, or a key distribution scheme for issuing symmetric keys). The system parameters of all the $|\mathbb{I}|$ credential schemes, denoted by Γ, are known to the initiators. Moreover, an initiator I may obtain credentials from some initiators (including itself). This means that I is also assigned with a unique identity $U \in \mathbb{U}$.

(2) **System self-forming**. The main functionality at this stage is to admit outside users. Suppose an outside user should be admitted according to the policy. Then, an initiator $I \in \mathbb{I}$ assigns the user with a unique identity $U \in \mathbb{U}$, and issues the user a credential cre_U. Both operations are done via the underlying authenticated (private) channel mentioned above. Besides, the initiator also sends Γ to the newly admitted outside user, who will only trust credentials that can be verified using Γ.

(3) **Optional initiators downgradation**. After the system has been formed, the initiators "downgrade" their roles to normal users. This is important because the initiators are the valuable nodes an adversary would like to target.

(4) **Establishment of security associations**. At any point in time, two admitted users can always establish a security association between them (e.g., a common secret). Notice that two users could establish a security association during the process of *system self-forming* and well before the process of *optional initiators downgradation*. The method of establishing a security association depends on the credential schemes employed by the initiators.

2.3.1 *Extensions.* We consider two types of extensions. One is to address the issue that more nodes need be admitted after the initiators downgraded their roles to become normal users. Note that this extension may effectively imply that the initiator set $\mathbb{I}_i$ is not necessarily fixed. The other is to "emulate" a threshold admission policy when it is necessary.

What if more nodes need be admitted after the optional initiators downgradation? For this purpose, we sketch two extensions.

(1) Suppose there are more than ℓ trusted initiators at the beginning, but we let only ℓ of them (we may call them "active initiators") act as initiators for admitting outside users. The other initiators may be called "inactive initiators," whose information (e.g., their public keys) is nevertheless disseminated by the active initiators to the admitted users. Suppose that, after the active initiators downgrade their roles, more users need be admitted. This can be fulfilled by activating the inactive initiators, whose (for example) public keys were already disseminated. Indeed, the need for admitting more users itself may be observed by the inactive initiators themselves (e.g., they can monitor how effective the MANET is).
Note that, due to the following security consideration, we may not let the downgraded initiators reactivate themselves as initiators. Specifically, an initiator that has admitted any user has exposed the fact "who is an initiator" and "who is more valuable from the adversary's perspective." This allows the adversary to *adaptively* select nodes to compromise for more significant damage. In contrast, the inactive initiators may not have been exposed to the adversary until they start to admit outside users.

(2) We observe that a MANET may be bootstrapped by some third party, which actually designated some initiators to establish the MANET in question. Although the third party stays offline, it may be able to observe the

aggregate state of the MANET (e.g., whether a targeted application has been accomplished). It would be reasonable to assume that the offline trusted third party is never compromised, and may know the cryptographic secrets of the initiators. Suppose that, after the initiators have downgraded their roles, more users need be admitted. Then, the offline third party can let new initiators join the MANET. Moreover, the new initiators may actually reuse the cryptographic keys of the past initiators (i.e., they are in a sense clones of the downgraded initiators); this would be more efficient than the above extension.

What if it is necessary to enforce a threshold admission policy? In the basic scheme explored above, we assumed that $\ell > s = 1$. Now we present an extension so as to emulate a threshold admission policy with $s > 1$. The extension is highlighted below.

—Whenever a user encounters an initiator, the user requests to be admitted unless it has been admitted s times by s initiators, respectively.

—Whenever two admitted users need to establish a security association, there are three cases.

(1) Both users have been admitted s times. In this case, the security association may be assigned with the highest trust by each other.

(2) One user has been admitted s times, but the other has been admitted $s' < s$ times. Then, a security association can also be established, but the trust put in each other may be asymmetric (e.g., the degree the former trusts the latter is less than the degree the latter trusts the former).

(3) Both users have not been admitted s times. Then, a security association can still be established, while less trust may be put in each other.

2.3.2 *Discussion: Toward a holistic security framework for MANETs.* The above framework does not address the revocation of compromised users, which is believed to be an orthogonal issue. This is mainly because, before a node is revoked, the fact that it has been compromised must be detected, say, by some intrusion detection systems. As a result, we focus on the worst case scenario where no such detection capability is assumed. Nevertheless, we hope that the bootstrapping framework can serve as a base so that, for example, the framework can be extended to a holistic framework for MANETs security. The holistic framework should include the detection and revocation of compromised users, and is left as a challenging open problem.

2.4 Adversarial Model

We consider a *hybrid* adversarial model, where the term hybrid means the following. From a computational (or cryptographic) perspective, we assume that the adversary is a probabilistic polynomial-time algorithm. This means that the adversary cannot break any cryptographic scheme that is proven secure in the modern cryptography framework. Moreover, the adversary cannot break the source identification and channel integrity of an assumedly authenticated channel; or, the adversary cannot break the source identification,

channel integrity, and channel confidentiality of an assumedly authenticated private channel.

From a system security perspective, we assume that the adversary may be able to capture some users who are within a short (e.g., one-hop) distance of it, and then extract their secrets (e.g., cryptographic keys) stored on their devices. For tamper-resistant hardware modules (if provided), we assume that it always takes some time for the adversary to compromise the module—starting from the point in time that the device is at the hand of the adversary. This means that the malicious and capable owner of a hardware module could always have compromised it, and that the resulting damage (e.g., compromise of data confidentiality due to the compromise of cryptographic keys) must be handled at a higher layer of attack-resilience management (e.g., downgradation of initiators' roles).

Specifically, we consider the following attacks.

(1) Adversary acting as an initiator: In the setting of distributed bootstrapping of MANETs, the initiators do not know in advance who are the users that will be admitted. Symmetrically, the users to be admitted into a MANET do not necessarily know in advance who are the initiators (or what credentials suffice someone as an initiator). Therefore, an adversary could claim to be an initiator and could admit users into their own MANETs. This is inevitable in the absence of a central authority. As a consequence, there could be multiple MANET overlays on top of the same physical networks, where each overlay corresponds to a (set of) initiator(s) and can be uniquely identified through its MANET-wide system parameters Γ. Note that, dependent upon a user's policy, a user may or may not join multiple MANETs. The goal of the adversary is to disrupt the establishment of security associations between users that are admitted into a MANET that is bootstrapped by some trusted initiators.

(2) Adversary compromising some legitimate users: Suppose an adversary can compromise α portion of the network users (possibly including the initiators), where $0 \leq \alpha \leq 1$. When the adversary successfully launches the Sybil attack by obtaining, for example, two identities with respect to the parameter set Γ of a MANET, we treat it as if the adversary compromised two users. To accommodate the worst case scenario, we assume that there is no good intrusion detection system that is deployed to identify the compromised users, and that all the compromised users are under the control of the adversary. In order to understand the impact of compromised users on the security associations between the noncompromised nodes, we classify adversarial behaviors into two dimensions: whether the adversary is able to compromise any initiators before the optional initiators downgradation, and whether an adversary is *random* or *adaptive* in choosing users to compromise.

(3) Adversary launching the Sybil attack [Douceur 2002]: This is a very significant threat. For example, this attack could simply ruin security of the whole system when the adversary manages to obtain a threshold

number of shares of the private key of the threshold cryptosystem [Zhou and Haas 1999; Yi and Kravets 2003], or obtain a threshold number of shares of the system-wide secret bivariate polynomial [Saxena et al. 2005]. As a consequence, the system-wide private key or the system-wide bivariate polynomial is compromised. There are two types of Sybil attack relevant to the setting of this article.

(a) Type I Sybil attack: An adversary may possess multiple out-of-band identities or credentials before being admitted into a MANET, where the out-of-band identities or credentials are relevant to the policy for admitting outside users. This is possible, for example, when a policy says "any one with a public key certificate issued by XYZ can be admitted into the network." This is because an adversary may have illegally obtained multiple such certificates. Countering Type I Sybil attacks is beyond the scope of the framework; nevertheless, a method detailed below and meant for countering Type II Sybil attacks would mitigate Type I Sybil attacks as a piggyback.
(b) Type II Sybil attack: An adversary with a single out-of-band identity or credential may obtain multiple identities with respect to a MANET. This is possible because any initiator may be allowed to admit users, and thus a dishonest outside user may get admitted multiple times with respect to different initiators. This type of Sybil attacks does not have a counterpart in existing threshold cryptosystem based bootstrapping schemes. Fortunately, as we will see in Sections 3-4, some practical methods could be exploited to alleviate not only Type II Sybil attacks, but also Type I Sybil attacks.

2.5 Properties

In order to understand and evaluate a bootstrapping scheme, we consider the following properties from a whole-system perspective: correctness, convenience, availability, robustness, efficiency, and security. Unlike traditional analysis of security-related properties of cryptographic primitives and protocols, we aim to understand the security-related properties from a whole system perspective. We believe that these properties are more relevant because we allow (possibly many) users to be compromised. Given that we are not aware of any well-accepted formalism for rigorously measuring or quantifying such system-oriented security properties, which we believe indeed are a challenging open problem, the presentation is necessarily informal at this stage of our knowledge.

—Correctness: Suppose that every participant is honest. Then this property says that, at any point in time, two admitted users can always establish a security association (e.g., a common secret).
—Convenience: By convenience, we mean under how many possibilities an outside user can be admitted into the network. This can be defined by ℓ/s, where ℓ is the number of initiators, and s is number of initiators that are necessary to admit a user (i.e., a user needs to get s approvals from s

initiators in order to get admitted). We notice that a larger ℓ/s implies a better convenience.[3]

—Availability: By availability, we mean under what circumstances the admission function will become not available. This should reflect the effort an adversary has to take in order to succeed in such a denial-of-service attack. This can be defined as $\ell - s + 1$, meaning that in order to make the system not available, the adversary has to impose denial-of-service attacks on $\ell - s + 1$ initiators. A larger $\ell - s + 1$ implies a better availability.

—Robustness: This property specifically concerns whether a scheme is subject to the aforementioned routing-security interdependence cycle problem or the proximity-caused insecurity problem. We say a bootstrapping scheme is robust if it is not subject to any of the two problems.

—Efficiency: We consider computational, communication, and storage complexities imposed on each individual user (i.e., an admitted user or an initiator) during and after the process of a user being admitted.

—Security: In a bootstrapping scheme, security aspects of interest are:

(1) Impersonation-tolerance. This property aims to capture the following: even if the adversary can act as initiators, MANETs initiated by honest initiators can still be securely bootstrapped. In other words, there is always a secure security association between two honest users, if and only if they were admitted into a MANET that was bootstrapped by some honest initiator(s).
More precisely, denote by U_i the identity issued to a user when the user joins its ith MANET, and $\Gamma_{U,i}$ the corresponding MANET parameters.[4] Furthermore, denote by $\Gamma_U = \{\Gamma_{U,1}, \ldots, \Gamma_{U,w_U}\}$ the set of parameters corresponding to the w_U MANETs user U joined, and by $\Gamma_V = \{\Gamma_{V,1}, \ldots, \Gamma_{V,w_V}\}$ the set of parameters corresponding to the w_v MANETs user V joined. Denote by Γ the system parameters of a MANET bootstrapped by some honest initiator(s). Then, impersonation-tolerance means that the two users can establish a security association if and only if $\Gamma \in \Gamma_U \cap \Gamma_V$.

(2) Intrusion-tolerance. Suppose an adversary can compromise α portion of the network users (possibly including the initiators), where $0 \leq \alpha \leq 1$. We need to investigate the impact of the compromised nodes on the security associations established between the noncompromised nodes. We notice that such an analysis accommodates both the case of the adversary being active (i.e., $0 < \alpha \leq 1$) and the case of the adversary being passive (i.e., $\alpha = 0$). We also observe that when a user U is compromised, the

[3]A more involved definition is possible so that we can differentiate, for example, the case of $s = \ell = 1$ from the case of $s = \ell > 1$.

[4]This is well defined because it is locally maintained by the user. Care must be taken in the case that the policy regulates $s > 1$, namely that a user may be admitted by multiple initiators corresponding to the same Γ. In this case, the identity assigned to an honest user the first time (by an initiator) will be used with respect to all the initiators corresponding to Γ.

security association between U and any other (even noncompromised) user is necessarily compromised.

(3) Sybil-tolerance. This captures the degree that the Sybil attack can be tolerated or mitigated. Ideally, we would like the degree to be precisely measured. Since it may not always be the case, we may only be able to argue, for example, that one scheme or configuration is more subject to the (Type I and/or Type II) Sybil attack. Intuitively, the better the more Sybil identities can be tolerated.

2.6 Parameter Selections

We explore how a practitioner may select system parameters for a bootstrapping scheme. Two crucial parameters are:

—When should the initiators downgrade their roles to normal users? Intuitively, the initiators should downgrade their roles after a MANET need not admit outside users anymore. In the case of multiple initiators, it would be ideal that the decision be made by the initiators in a distributed fashion (i.e., without relying on any coordination).

—How many initiators are needed? Intuitively, the more initiators the better. However, this is unrealistic because there may not be so many different types of tamper-resistant hardware modules (for better diversity and security), or because there is a financial issue that indicates that only a small number of tamper-resistant hardware modules are available.

2.7 Discussion on the Usefulness and Generality of the Framework

2.7.1 *On the usefulness of the framework.* On one hand, the framework provides a common base for evaluating and comparing relevant bootstrapping schemes. This is necessary for deepening our understanding of the bootstrapping problem, and serves as a solid step towards identifying the optimal solution. On the other hand, as we will see in Sections 3-4, the framework leads to two (classes of) instantiations that, unlike existing schemes, avoid both the "routing-security interdependence cycle" problem and the "proximity-caused insecurity" problem.

2.7.2 *On the generality of the framework.* The framework can accommodate many bootstrapping schemes, including the two new (classes of) schemes detailed in Sections 3-4. To see this, we consider the schemes that are dominating in the literature, where the credential schemes chosen by the initiators are based on threshold public key cryptosystems.

Suppose the credential schemes of the initiators are based on threshold public key cryptosystems. On one hand, if we fix the initiator set to be $\mathbb{I} = \{I_1, \ldots, I_\ell\}$ and the threshold s to be $1 < s \leq \ell$ for the whole life-time of the system, we immediately obtain the schemes of Zhou and Haas [1999], and Yi and Kravets [2003]. Furthermore, by coupling the fixed $\mathbb{I}$ and s with the assumption made in Bechler et al. [2004], and Xu and Iftode [2004] (i.e., the network possesses a certain structure such as cluster-based or locality-based),

we immediately obtain their schemes. On the other hand, if we allow the initiator sets $\mathbb{I}_0, \mathbb{I}_1, \ldots$ to be dynamic in the sense that newly admitted nodes are also allowed to admit outside users, we immediately obtain the scheme of Kong et al. [2001].

3. DISTRIBUTED AND SECURE BOOTSTRAPPING BASED ON PUBLIC KEY CRYPTOSYSTEMS

In this section we present a concrete instantiation of the above framework, assuming that the credential schemes of the initiators are based on public key cryptosystems. Note that, this does not necessarily mean that the authentication of outside users is also based on public key cryptosystems, because authentication, or more precisely the enforcement of the policy, is independent of the effort in the present article. Furthermore, we assume that there are a *fixed* number of initiators, meaning $\ell = \ell_0 = \ell_1 = \ldots$ and $\mathbb{I} = \mathbb{I}_0 = \mathbb{I}_1 = \ldots = \{I_1, \ldots, I_\ell\}$, and that any initiator is authorized to admit outside users, meaning $t = t_0 = t_1 = \ldots = 1$. We will discuss how such a scheme can be extended to mimic a threshold admission policy.

3.1 Construction

In this scheme, the initiators simply issue cryptographic credentials using standard digital signature schemes. We do not pin down any concrete signature scheme, because any such scheme could be deployed as long as it satisfies the standard security definition [Goldwasser et al. 1988]. Practical constructions can be found in, for example, Rivest et al. [1978], and Bellare and Rogaway [1996].

(1) **Setup of initiators**. Each initiator $I_i \in \mathbb{I}$, $1 \leq i \leq \ell$, establishes a public key cryptosystem, corresponding to which the pair of public and private keys $(\mathsf{PK}_i, \mathsf{SK}_i)$ are called an authority key. Suppose $\Gamma = \{\mathsf{PK}_1, \ldots, \mathsf{PK}_\ell\}$ are known to all the initiators, but are not necessarily known to the outside users at this point in time—explaining why this is different from standard public key infrastructures. Each initiator I_i, $1 \leq i \leq \ell$, also chooses a pair of public and private keys (pk_U, sk_U) for acting as a normal user U, where pk_U is certified by I_i's own authority private key SK_i (but can also be certified by other initiators, if needed).

(2) **System self-forming**. Suppose a yet-to-be-admitted outside user moves within a certain distance of an initiator $I_i \in \mathbb{I}$. If the outside user should be admitted according to the policy (enforcement of the policy is conducted over the underlying authenticated channel), I_i assigns the user a unique identity U and certifying U's public key pk_U via the underlying authenticated channel, where pk_U may be chosen by U on the fly. (Such a process may require the user to prove, via a simple cryptographic means, that it does know sk_U.) In addition, I_i sends U information including $\Gamma = \{\mathsf{PK}_1, \ldots, \mathsf{PK}_\ell\}$; this means that U will only accept as valid credentials that can be verified using some $\mathsf{PK}_j \in \Gamma$.

Properties	Centralized Scheme ($\ell = s = 1$)	Our Scheme ($\ell > s = 1$)	Threshold Scheme ($\ell \geq s > 1$)
convenience	no	best	good
availability	single point of failure	best	good
robustness	yes	yes	no
efficiency	practical	practical	less practical
impersonation– tolerance	yes	yes	yes
intrusion–tolerance	no	strong	weak
Sybil–tolerance	best	good	weak

Fig. 1. Comparison of secure bootstrapping schemes based on public key cryptosystems.

(3) **Optional initiators downgradation**. Suppose an initiator I_i decides to downgrade its role to a normal user, then it securely erases its authority key SK_i (but keeps its normal private key sk_i). Exactly when this should take place depends on certain system parameters; see Section 5 for details. Ideally, such a decision is made in a distributed fashion (without relying on any coordination).

(4) **Establishment of security associations**. At any point in time two admitted users with certified (pk_U, sk_U) and (pk_V, sk_V), respectively, can always authenticate each other's public key and thus establish a security association between them. There are a large family of protocols for this purpose (cf., for example, Krawczyk [2003]). Notice that two users could establish a security association during the process *system self-forming* and well before the process of *optional initiators downgradation*.

3.2 Analysis

The correctness of the scheme is clear because any two admitted users hold the same list of authority public keys Γ. Specifically, suppose that an outside user U is admitted by initiator I_i, and another outside user V is admitted by initiator I_j. Then, U holds $\Gamma = \{\mathsf{PK}_1, \ldots, \mathsf{PK}_\ell\}$ and pk_U is certified by SK_i, and V holds Γ and pk_V is certified by SK_j. Since U accepts pk_V as a valid credential and V accepts pk_U as a valid credential, they can immediately establish a security association.

To see the advantages of the newly proposed bootstrapping schemes, we compare it with a centralized scheme and a threshold cryptosystem based scheme, respectively. The comparison is from the perspectives of convenience, availability, robustness, efficiency, and security. The comparison results are highlighted in Figure 1, and elaborated below.

3.2.1 *Centralized Scheme vs. Our Distributed Scheme.* Notice that in a centralized scheme we have $\ell = s = 1$, and in our distributed scheme we have $\ell > s = 1$. For compatibility, we assume that in the centralized case, a user is only admitted when moving into a short (e.g., one-hop) distance of the initiator.

Convenience. Compared with the centralized scheme, our distributed scheme achieves better convenience because $\ell > 1$. This means that an outside user

can be admitted by any initiator (out of the ℓ initiators) it encounters, rather than has to be admitted by the single designated initiator. It would take a much longer time for an outside user to be admitted in the centralized scheme.

Availability. Compared with the centralized scheme, our distributed scheme clearly achieves better availability because $\ell > 1$, meaning that there is no single point of failure. In the centralized scheme, the adversary could capture the authority and takes its out of the system before the system is formed.

Robustness. Both the centralized scheme with $\ell = s = 1$ and our distributed scheme with $\ell > s = 1$ can avoid both the routing-security interdependence cycle problem and the proximity-caused insecurity problem.

Efficiency. We first consider computational cost. From the perspective of an admitted user, both the centralized scheme and our distributed scheme have the same complexity. This is because whenever a user needs to verify a certificate it had not encountered before, it needs to verify the initiator's signature anyway (i.e., it does not matter which initiator certified it). From the perspective of the initiators, our distributed scheme is more efficient because the cost of certifying or admitting the same number of outside users is amortized to ℓ initiators.

In terms of communication cost, our distributed scheme is slightly more expensive than the centralized one. This is because in our distributed scheme the admitting initiator needs to send every admitted user all the ℓ public keys $\Gamma = \{\mathsf{PK}_1, \ldots, \mathsf{PK}_\ell\}$. Whereas, in the centralized scheme $\ell = 1$. This should not be seen as a problem, even if the authenticated channel has a low bandwidth. This is because the authenticated channel can be used to transfer the hash of the Γ, which is then transferred using the normal wireless channel.

Now we consider storage complexity. From the perspective of an admitted user, our distributed scheme imposes a complexity that is ℓ times of the one imposed by the centralized scheme. From the perspective of an initiator, our distributed scheme is more efficient because each initiator at most needs to keep the public keys it has certified.

Security. We detail security analysis below.

(1) Impersonation-tolerance. Suppose $\Gamma = \{\mathsf{PK}_1, \ldots, \mathsf{PK}_\ell\}$ is the parameter set corresponding to a MANET bootstrapped by some honest initiators. First, we show that *if* two honest users were admitted into the the same MANET bootstrapped by some honest initiators, then they can establish a security association. This is because the same Γ is used by the two users, and then two honest users can authenticate each other's public key.
Second, we show that only if two honest users were admitted into the the same MANET bootstrapped by some honest initiators, they can establish a security association. Suppose U is admitted by some initiator I_i according to the policy, and V is admitted by a fake initiator. Then, U holds Γ and pk_U is certified by SK_i, and V holds Γ' and pk_V is certified by the private key of the impersonator. There are two cases. In the case $\Gamma = \Gamma'$, it must hold that the adversary can, with a nonnegligible probability, issue a

digital signature to V with respect to some PK_j for some $1 \leq j \leq \ell$. Since the corresponding SK_j was not compromised by the adversary, this immediately leads to that the digital signature scheme corresponding to PK_j is not secure, which contradicts with the assumption. In the case $\Gamma \neq \Gamma'$, then U simply rejects V according to the protocol, and thus will not establish a security association with V. Note that this discussion holds even if $\Gamma \cap \Gamma' \neq \emptyset$, as long as $\Gamma \neq \Gamma'$.

(2) Intrusion-tolerance. Suppose α portion of the network users are compromised. Recall that we classify adversarial behaviors according to two dimensions. One is about whether the adversary is able to compromise any initiators before the *optional initiator downgradation*. The other is about whether an adversary is *random* or *adaptive* in choosing users to compromise. Therefore, there are four cases.

   (a) A random adversary is unable to compromise any initiator before the optional initiators downgradation. In the centralized case, the adversary cannot compromise the communication between any pair of noncompromised users. The same is true in our distributed scheme.
   (b) An adaptive adversary is unable to compromise any initiators before the optional initiators downgradation. In the centralized case, the adversary cannot compromise the security association between any pair of noncompromised users. The same is true in our distributed scheme.
   (c) A random adversary is able to compromise some initiator(s) before the optional initiators downgradation. In the centralized case, the whole system is compromised in the sense that the adversary can arbitrarily admit any users into the network. Therefore, all the communications are compromised, except those between two noncompromised users (there are such users because they may be admitted into the network before the initiator is compromised).
   In our distributed scheme, it would be fair to consider the scenario that exactly one of the initiators is compromised. Given that, our distributed scheme ensures that at least the same number of honest outside users will be admitted into the network. This means that it would be easier for a later detection of the compromise of the initiator (as well as the users admitted by the compromised initiators). Full exploration towards this end is beyond the scope of the present article; however, our distributed scheme does provide the base upon which other systems (e.g., intrusion detection systems) can build.
   (d) An adaptive adversary is able to compromise some initiator(s) before the optional initiators downgradation. This is the same as in the above case that "a random adversary is able to compromise some initiators before the optional initiators downgradation."

(3) Sybil-tolerance. A successful Sybil attack could allow a dishonest outside user to be admitted multiple times (i.e., the Type II Sybil attack specified in the adversarial model). In the centralized case, the system is less subject to the Type II Sybil attack if the central authority can recognize two

admission requests from the same outside user.[5] In the distributed case, if we assume the same capability on the initiators, an adversary can at most be admitted ℓ times. Since $n >> \ell$, the Sybil attack is mitigated. Moreover, the Sybil attack may be detected by letting the initiators exchange information about the admitted users.

The above analysis shows that our distributed scheme is advantageous in terms of security when the adversary is able to break a single initiator before the initiator downgrades its role to a normal user. Notice that it takes, on average, approximately ℓ times longer in time for the initiator in the centralized case to downgrade its role than in our distributed scheme (See Section 5 for details); this explains why it somewhat suffers from the security-routing interdependence cycle problem. A more comprehensive comparison is presented in Figure 1.

3.2.2 *Threshold Cryptosystem-based Scheme vs. Our Distributed Scheme.* Notice that in our distributed scheme $\ell > s = 1$, and in a threshold cryptosystem based scheme $\ell \geq s > 1$.

Convenience. Compared with a threshold public key cryptosystem based scheme, our distributed scheme achieves better convenience because $s = 1$. This is because an outside user can be admitted by any initiator (out of the ℓ initiators) it encounters, rather than has to be admitted by $s > 1$ initiators.

Availability. Since $\ell - 1 > \ell - s$ when $s > 1$, our distributed scheme clearly achieves a better availability.

Robustness. Our distributed scheme with $\ell > s = 1$ does avoid both the routing-security interdependence cycle problem and the proximity-caused insecurity problem. However, as mentioned in the Introduction, the threshold scheme suffers from at least one of the two problems.

Efficiency. Let's first consider computational cost. From the perspective of an initiator, our distributed scheme is clearly advantageous over a threshold cryptosystem-based one. This is because each initiator only needs to be involved in admitting some of the admitted users. From the perspective of an outside user, our distributed scheme is at least as efficient as a threshold cryptosystem-based one.

Regarding communication complexity, our distributed scheme is much more efficient because in threshold cryptosystem-based cases at least s initiators must be involved in order for an outside user to be admitted.

Finally we examine storage complexity. From the perspective of an initiator, our distributed scheme is more efficient as each initiator only needs to be aware of some of the outside users. From the perspective of an outside user, our distributed scheme is at least as efficient as a threshold cryptosystem-based scheme.

[5]For example, an initiator may record the driver's license numbers of the admitted users. Provided that driver's licenses are not faked, the initiator can always make sure that a user being admitted was not admitted before—simply by making sure that the photo matches the person. This actually helps alleviate Type I Sybil attacks specified in the adversarial model.

Security. In terms of security, we have the following:

(1) Impersonation-tolerance. In a fashion similar to the above reasoning that our distributed scheme is impersonation-tolerant, we can show that a threshold cryptosystem based scheme is also impersonation-tolerant.
(2) Intrusion-tolerance. Suppose α portion of the network users are compromised. Again, there are four subcases.
   (a) A random adversary is unable to compromise any initiator(s) before the optional initiators downgradation. In both schemes, the adversary cannot compromise the security association between any pair of noncompromised users.
   (b) An adaptive adversary is unable to compromise any initiator(s) before the optional initiators downgradation. In both schemes, the adversary cannot compromise the security association between any pair of noncompromised users.
   (c) A random adversary is able to compromise some initiator(s) before the optional initiators downgradation. In both schemes, compromise of a certain threshold number of initiators allows the adversary to admit dishonest users. Specifically, before the threshold is reached, our distributed scheme allows the adversary to admit perhaps more dishonest users than a general threshold cryptosystem does; after the threshold is reached (but before all initiators are compromised), our distributed scheme can admit more honest users than a general threshold cryptosystem does—this would make it easier for the later detection of compromise of the initiator (as well as the users admitted by the compromised initiators).
   (d) An adaptive adversary is able to compromise some initiator(s) before the optional initiators downgradation. This is the same as in the above case that "a random adversary is able to compromise some initiators before the optional initiators downgradation."
(3) Sybil-tolerance. In the general threshold cryptosystem case where the threshold $1 < s \leq \ell$, the system is subject to Type I and Type II Sybil attacks specified in the adversarial model. This is because the admission requests will necessarily come from some remote users (e.g., not within a visual distance), and possibly because $\ell \geq 2s$. In our scheme (i.e., $s = 1$), if we assume that the initiators can recognize the users they have admitted, an adversary can at most be admitted ℓ times, regardless how many out-of-band credentials a Type I Sybil attacker possesses. Since $n >> \ell$, the Type II Sybil attack is mitigated (whereas Type I Sybil attack is naturally avoided as long as, for example, driver's licenses are not faked). Moreover, the Type II Sybil attack in our distributed scheme may be detected by letting the initiators exchange information about the admitted users. This of course will incur some significant communication complexity.

The above analysis shows that our scheme is advantageous, especially because it can avoid both the routing-security interdependence cycle problem and the proximity-caused insecurity problem. The intrusion-tolerance of the

threshold scheme may be deceptively better than our scheme, while actually it is not. This is because an adversary can compromise the whole network by compromising a threshold number of any nodes within the whole lifetime of the network. Things are actually much worse, because it is subject to both Type I and Type II Sybil attacks. In particular it is not clear how Type I Sybil attacks can be alleviated without suffering from the proximity-caused insecurity problem. The comparison is highlighted in Figure 1.

3.2.3 *Summary.* The above analysis shows, among other things, that our distributed scheme has advantages over threshold cryptosystems-based scheme, especially in terms of robustness (i.e., avoiding both the routing-security interdependence cycle problem and the proximity-caused insecurity problem), intrusion-tolerance, and Sybil-tolerance.

3.3 Extension

On one hand, the extension of the framework (see Section 2.3), namely the one for accommodating that more users need be admitted after the initiators downgrade their roles to normal users, can be naturally inherited in this scheme. Note that this extension may effectively imply that the initiator set $\mathbb{I}_i$ is not necessarily fixed; this is because, for example, only a smaller number of initiators may be activated for admitting more users. Thus, here we elaborate on the other extension of the basic scheme to enforcing a threshold admission policy.

Recall that in the basic scheme explored above, we assumed that $\ell > s = 1$. Now we present an extension so as to emulate a threshold cryptosystem-based scheme with $s > 1$. This is useful when threshold admission is absolutely necessary. The idea is highlighted below.

—An initiator is admitted by s (e.g., randomly picked) initiators.
—Whenever a user encounters an initiator, the user requests to be admitted unless it has been admitted s times by s initiators, where s is affiliated with Γ (or part of Γ). It is important to note that a user will only need to have a single pair of public and private keys, but the public key are certified by s initiators with respective certificates.
—Whenever two admitted users need to establish a security association, there are three cases.

(1) Both users have been admitted s times. In this case, the security association may be assigned with the highest trust by each other.
(2) One user has been admitted s times, but the other has been admitted $s' < s$ times. Then, a security association can also be established, but the trust put in each other may be asymmetric (e.g., the degree the former trusts the latter less than the degree the latter trusts the former).
(3) Both users have not been admitted s times. Then, a security association can still be established, while less (and possibly also asymmetric) trust may be put in each other.

Note that the above virtual threshold admission scheme has the desired robustness (i.e., avoiding both the routing-security interdependence cycle problem and the proximity-caused insecurity problem). It is interesting to note that the virtual threshold admission scheme brings a flexible scale of "fine-grained trust" between the admitted users.

4. DISTRIBUTED AND SECURE BOOTSTRAPPING BASED ON SYMMETRIC KEY CRYPTOSYSTEMS

In this section we present a class of instantiations of the framework, assuming the credential schemes of the initiators are based on symmetric key cryptosystems. Similarly, we also assume that there are a *fixed* number of initiators, meaning $\ell = \ell_0 = \ell_1 = \ldots$ and $\mathbb{I} = \mathbb{I}_0 = \mathbb{I}_1 = \ldots = \{I_1, \ldots, I_\ell\}$, and that any single initiator is authorized to admit outside users, meaning $t = t_0 = t_1 = \ldots = 1$. We will discuss extensions where $\mathbb{I}_i$ may not be fixed (e.g., when more users need be admitted after the initiators downgrade their roles to normal users).

4.1 Construction

To be concrete, we adopt the basic scheme of Blom [1984]. The basic idea underlying this scheme is the following: Suppose a key distribution server picks a random bivariate polynomial of degree d, $f(x, y) = \sum_{i,j=0}^{d} c_{ij} x^i y^j$, over a finite field $\mathbb{F}_q$, where q is a large prime (e.g., $|q| = 128$). The polynomial has the property that $f(x, y) = f(y, x)$ for all $x, y \in \mathbb{F}_q$. Suppose that each user has a unique identity $u \in \mathbb{F}_q$, and given a polynomial $f(u, y)$. Clearly, two users of identities u and v can immediately derive a common secret $f(u, v) = f(v, u)$.

Before we present the details of our scheme, we discuss its basic ideas. Recall that there are ℓ initiators that need to admit n outside users. Suppose the expected/tolerated portion of compromised nodes is α, where $0 \leq \alpha < 1$. We let the initiators play the role of the key distribution servers in the basic scheme of Blom [1984], which is called a building-block key distribution scheme. Specifically,

(1) Let each pair of initiators (I_i, I_j), where $1 \leq i, j \leq \ell$ and possibly $i = j$, share a secret bivariate random polynomial $f_{ij} = f_{ji}$ of degree d such that $f_{ij}(x, y) = f_{ij}(y, x) = f_{ji}(x, y) = f_{ji}(y, x)$. Therefore, each initiator keeps ℓ polynomials.

(2) Partition the identity space into subspaces, each of which will be administered by an initiator. For example, we can partition $\mathbb{F}_q$ into ℓ identity subspaces with each of size $\lfloor q/\ell \rfloor$; this does not incur any problem because $q >> \ell$ (e..g, $|q| = 128$, n is at the order of 10^3, and ℓ is at the order of 10 or 10^2). Denote by ID_i the identity subspace administered by initiator I_i, where $1 \leq i \leq \ell$. When an outside user is admitted by initiator I_i, it is assigned with a unique identity $U \in ID_i$ and given polynomials $f_{i1}(U, \cdot), \ldots, f_{i\ell}(U, \cdot)$.

As a result, two admitted users, U admitted by I_i and V admitted by I_j, where possibly $i = j$, can immediately establish a security association since $f_{ij}(U, V) = f_{ji}(V, U)$.

Now we present the details.

(1) **Setup of initiators**. Each pair of initiators (I_i, I_j), where $1 \leq i, j \leq \ell$, establishes an appropriate symmetric bivariate polynomial $f_{ij} = f_{ji}$ of degree d such that $f_{ij}(x, y) = f_{ij}(y, x) = f_{ji}(x, y) = f_{ji}(y, x)$, where d will be determined later (see security analysis below). Each initiator I_i selects a unique identity $U \in ID_i$ (an arbitrary one suffices, but a counter-like one would simplify maintenance), and assigns polynomials $f_{i1}(U), \ldots, f_{i\ell}(U)$ to itself.

(2) **System self-forming**. Suppose a yet-to-be-admitted outside user moves to within a certain distance of an initiator I_i. If the outside user should be admitted according to the policy, then I_i selects for the outside user a unique identity $V \in ID_i$, sends it the polynomials $f_{i1}(V, \cdot), \ldots, f_{i\ell}(V, \cdot)$ via the underlying authenticated private channel. (As said before, if it is not practical to transfer ℓ polynomials of degree d over this channel because of its limited bandwidth, I_i can instead use this channel to send a temporary secret key to V. Then, this temporary key is used to protect the transfer of the ℓ polynomials over the normal wireless channel, and securely erased by both I_i and V afterwards.)

(3) **Optional initiators downgradation**. Suppose an initiator I_i with a normal user identity U decides to downgrade its roles to a normal user (see Section 5 on when such a decision is made), it securely erases the $f_{i1}(\cdot, \cdot), \ldots, f_{i\ell}(\cdot, \cdot)$'s but keeps $f_{i1}(U, \cdot), \ldots, f_{i\ell}(U, \cdot)$.

(4) **Establishment of security associations**. At any point in time, user U admitted by I_i and user V admitted by I_j can always establish a security association between them. This is because U has $f_{ij}(U, \cdot)$ and V has $f_{ji}(V, \cdot)$ such as $f_{ij}(U, V) = f_{ji}(V, U)$. Notice that two users could establish a security association during the process of system self-forming and well before the process of optional initiators downgradation.

4.2 Analysis

We compare our distributed scheme (i.e., $\ell > 1$) with an imagined centralized scheme (i.e., $\ell = 1$), and with the distributed scheme of Saxena et al. [2005]. Notice that Saxena et al. [2005] bears much similarity with the centralized scheme, except (1) the central authority can immediately disappear after admitting a threshold number, s, of outside users, and (2) any s admitted users can collaboratively admit other users. Therefore, the following discussions mainly focus on the centralized case. We highlight the comparison of the three schemes in 2.

Correctness. In our distributed scheme, any $U \in ID_i$ and any $V \in ID_j$ can establish a security association because they possess a common secret $f_{ij}(U, V) = f_{ji}(V, U)$. This property is also true in both the centralized scheme and the scheme of Saxena et al. [2005].

Convenience. Compared with a centralized scheme, our distributed scheme achieves better convenience because $\ell > 1$. This is because an outside user can be admitted by any initiator (out of the ℓ initiators) it encounters, rather

Properties	Centralized Scheme ($\ell = s = 1$)	Our Scheme ($\ell > s = 1$)	[Saxena et al. 2005] ($s > 1$)
convenience	no	good	good
availability	single point of failure	good	good
robustness	yes	yes	no
efficiency	best	good	good
impersonation– tolerance	yes	yes	yes
intrusion–tolerance	no	strong	weak
Sybil–tolerance	best	good	weak

Fig. 2. Comparison of secure bootstrapping schemes based on symmetric key cryptosystems.

than has to be admitted by the single designated initiator. We notice that the scheme of Saxena et al. [2005] is also convenient.

Availability. Compared with a centralized key management scheme, our distributed scheme clearly achieves better availability since $\ell > 1$. This means that there is no single point-of-failure in the presence of denial-of-service attacks. In the centralized case, the adversary could capture the authority and takes its out of the system–even if the adversary may not be able to compromise the authority's secrets. The scheme of Saxena et al. [2005] is highly available.

Robustness. Both the centralized and our distributed scheme can avoid the routing-security interdependence cycle problem and the proximity-caused insecurity problem. As mentioned in the Introduction, the scheme of Saxena et al. [2005] is subject to at least one of the two problems.

Efficiency and security. Since security of this type of schemes is closely related to its efficiency, we analyze them together. Recall that an adversary is allowed to compromise α portion of the nodes (including possibly the initiators). It was shown in Blom [1984], and Blundo et al. [1992] that an adversary compromising at most d users, denoted by Δ', cannot derive any information (in an information-theoretic sense) about $f(U, V)$ for $U, V \notin \Delta'$.

(1) Impersonation-tolerance. First, we show that *if* two honest users were admitted into the the same MANET bootstrapped by two honest initiators, i and j, respectively, then they can establish a security association. Suppose the users are assigned identities U and V, respectively. Then, a common secret is given by $f_{ij}(U, V) = f_{ji}(V, U)$.
Second, we show that *only if* two honest users were admitted into the same MANET bootstrapped by some honest initiators, then they can establish a security association. Suppose U is admitted by some initiator I_i according to the policy, and V is admitted by a fake initiator. Then, U holds $f_{i1}(U, \cdot), \ldots, f_{i\ell}(U, \cdot)$, where $f_{ij}(U, V)$ is indistinguishable from a random string (when the credential issuance channel is an authenticated private one), or is indistinguishable from a pseudorandom string (when the authenticated private channel is based on an authenticated channel). In either case, there is no common secret between U and V, except for a negligible probability. Note that in practice, whether two users have a common secret would have to be figured out after some interaction (e.g.,

a challenge-response interaction based on a message authentication code). As long as the message authentication code is secure in the modern cryptographic framework, V cannot cheat U into accepting that V is an admitted user.

(2) Intrusion-tolerance. Suppose α portion of the nodes may get compromised where $0 \leq \alpha < 1$. Again, there are four cases.
   (a) A random adversary is unable to compromise any initiator before the optional initiators downgradation. In the centralized case, we can set the degree of the bivariate polynomial d to be $\lceil \alpha(n+\ell) \rceil$, which means that each outside user's storage complexity is $\lceil \alpha(n+\ell) \rceil + 1$ elements (e.g., each with length typically 128 bits). Each initiator's storage complexities before and after the optional initiators downgradation are $(\lceil \alpha(n+\ell) \rceil)^2 + \lceil \alpha(n+\ell) \rceil + 2$ elements and $\lceil \alpha(n+\ell) \rceil + 1$ elements, respectively. The communication complexity between an outside user and an initiator is $\lceil \alpha(n+\ell) \rceil + 1$ elements. The above discussion also applies to the scheme of Saxena et al. [2005].
   In our distributed scheme, we assume on average that each initiator will admit the same number of outside users, namely $\lceil n/\ell \rceil + 1$, of which at most $\lceil \alpha(\lceil n/\ell \rceil + 1) \rceil$ may be compromised. Therefore, we can set each polynomial to be of degree $\lceil \alpha(\lceil n/\ell \rceil + 1) \rceil$. This means that each outside user's storage complexity is $(\lceil \alpha(\lceil n/\ell \rceil + 1) \rceil + 1)\ell$. Each initiator's storage complexities before and after the optional initiators downgradation are $((\lceil \alpha(\lceil n/\ell \rceil + 1) \rceil)^2 + \lceil \alpha(\lceil n/\ell \rceil + 1) \rceil + 2)\ell$ elements and $(\lceil \alpha(\lceil n/\ell \rceil + 1) \rceil + 1)\ell$ elements, respectively. The communication complexity between an outside user and an initiator is $(\lceil \alpha(\lceil n/\ell \rceil + 1) \rceil + 1)\ell$ elements.
   In summary, in both cases an adversary able to compromise at most $\lceil \alpha(n+\ell) \rceil$ users is unable to compromise any communication between two non-compromised users. Compared with the centralized scheme, the initiator storage complexity in our distributed scheme is $O\left(\lceil \alpha(n+\ell) \rceil)^2 \cdot \sqrt{\ell}\right)$ more expensive, and an admitted user's (including the case the user is downgraded from an initiator) storage complexity in our distributed scheme is about ℓ more elements.
   (b) An adaptive adversary is unable to compromise any initiator before the optional initiators downgradation. In the centralized case, the adversary is able to compromise at most $\lceil \alpha(n+\ell) \rceil$ users, perhaps after the optional initiators downgradation. Therefore, we set the degree of the bivariate polynomial to $\lceil \alpha(n+\ell) \rceil$, which means that each outside user's storage complexity is $\lceil \alpha(n+\ell) \rceil + 1$ elements. Each initiator's storage complexities before and after the optional initiators downgradation are $(\lceil \alpha(n+\ell) \rceil)^2 + \lceil \alpha(n+\ell) \rceil + 2$ elements and $\lceil \alpha(n+\ell) \rceil + 1$ elements, respectively. The communication complexity between an outside user and an initiator is $\lceil \alpha(n+\ell) \rceil + 1$.
   In our distributed scheme, we still assume on average that each initiator will admit the same number of outside users, namely $\lceil n/\ell \rceil + 1$. Unlike in the random adversary case, we cannot assume that $\lceil \alpha(\lceil n/\ell \rceil + 1) \rceil$ of the $\lceil n/\ell \rceil + 1$ may be compromised as the adversary is adaptive.

Therefore, we let each polynomial be of degree $d = \lceil n/\ell \rceil$. This means that each outside user's storage complexity is $(\lceil n/\ell \rceil + 1)\ell$ elements. Each initiator's storage complexities before and after the optional initiators downgradation are $((\lceil n/\ell \rceil)^2 + 1)\ell + n + \ell$ elements and $n + \ell$ elements, respectively. The communication complexity between an outside user and an initiator is $(\lceil n/\ell \rceil + 1)\ell$ elements.

In summary, in both cases an adversary able to compromise at most $\lceil \alpha(n + \ell) \rceil$ users is unable to compromise any communication between two non-compromised users. The admitted user storage complexity in the recommended distributed scheme is about $\lfloor (1 - \alpha)(n + \ell) \rfloor$ more elements. The initiator complexity is $(\lceil n/\ell \rceil^2 + 1)\ell + n + \ell - \lceil \alpha(n + \ell) \rceil (\lceil \alpha(n + \ell) \rceil + 1) - 2$ more elements of $\log q$ bits.

(c) A random adversary is able to compromise some initiator(s) before the optional initiators downgradation. In the centralized case, the whole system is compromised.

In our distributed scheme, since we can assume on average that each initiator will admit about the same number of outside users, namely $\lceil n/\ell \rceil$, of which $\lceil \alpha(\lceil n/\ell \rceil + 1) \rceil$ may be compromised. Therefore, we can set each polynomial to be of degree $\lceil \alpha(\lceil n/\ell \rceil + 1) \rceil$. This means that each outside user's storage complexity is $(\lceil \alpha(\lceil n/\ell \rceil + 1) \rceil + 1)\ell$. Each initiator's storage complexities before and after the optional initiators downgradation are $(\lceil \alpha(\lceil n/\ell \rceil + 1) \rceil)^2 + \lceil \alpha(\lceil n/\ell \rceil + 1) \rceil + 2)\ell$ elements and $(\lceil \alpha(\lceil n/\ell \rceil + 1) \rceil + 1)\ell$ elements, respectively. The communication complexity between an outside user and an initiator is $(\lceil \alpha(\lceil n/\ell \rceil + 1) \rceil + 1)\ell$ elements.

In our distributed scheme, suppose U is admitted by I_i and V is admitted by I_j, where possibly $i = j$. Furthermore, suppose both U and V are not compromised. Then, the security association between U and V is not compromised, unless either I_i or I_j is compromised. In general, suppose the adversary compromises $a < \ell$ initiators, denoted by Δ, then the security association between any non-compromised U and V is not compromised, provided that $I_i \notin \Delta$ and $I_j \notin \Delta$, I_i admitted U and I_j admitted V.

In summary, in the centralized scheme, the whole system is compromised. Whereas, in our distributed scheme, there are $\lfloor (1 - \alpha)(\lceil n/\ell \rceil + 1)(\ell - a) \rfloor$ users they can still conduct secure communications.

(d) An adaptive adversary is able to compromise some initiator(s) before the optional initiators downgradation. In the centralized case, the whole system is compromised.

In our distributed scheme, we may still assume on average that each initiator will admit roughly the same number of outside users, namely $\lceil n/\ell \rceil$. Unlike the random adversary case, we cannot assume that $\lceil \alpha(\lceil n/\ell \rceil + 1) \rceil$ of the $\lceil n/\ell \rceil$ may be compromised as the adversary is adaptive. Therefore, we let each polynomial have degree $d = \lceil n/\ell \rceil$. This means that each outside user's storage complexity is $(\lceil n/\ell \rceil + 1)\ell$. Each initiator's storage complexities before and after the optional initiators downgradation are $((\lceil n/\ell \rceil)^2 + n + 2)\ell$ elements and $n + \ell$ elements,

respectively. The communication complexity between an outside user and an initiator is $(\lceil n/\ell \rceil + 1)\ell$ elements.
In summary, in the centralized case, the whole system is compromised. Whereas, in our distributed scheme, there are about $n - (\lceil n/\ell \rceil + 1) - \alpha n$ users they can still conduct secure communications.

(3) Sybil-tolerance. In the centralized case, the system may be less subject to Type II Sybil attacks provided that the centralized authority can recognize two admission requests from the same outside user (see Section 3 for how this may be fulfilled). In our distributed scheme, if we assume the same capability of an initiator, an adversary can at most be admitted ℓ times. We notice that $n >> \ell$. Moreover, Type II Sybil attacks may be detected by letting the initiators exchange information about the admitted users; this, of course, will incur extra communication complexity. Note that both the centralized scheme and our distributed scheme are not subject to Type I attacks. The scheme of Saxena et al. [2005] is arguably subject to Sybil attacks.

Note that the intrusion-tolerance of Saxena et al. [2005] is weak because the compromise of a threshold number of any nodes leads to the compromise of the whole system. This explains why the scheme of Saxena et al. [2005] was suggested for short-living MANETs. Finally, we note that when comparing our distributed scheme based on public key cryptosystems and our distributed scheme based on symmetric key cryptosystems, we recommend the former when storage and communication complexities are more important.

4.3 Extensions

On one hand, the extension of the framework (see Section 2.3), namely the one for accommodating the need for admitting more users after the initiators downgrade their roles to normal users, can be naturally inherited in this scheme. Thus, here we elaborate on the other extensions of the basic scheme.

As in the case of public key cryptosystems based credentials, the above scheme can also be configured to fulfill virtual threshold cryptosystems while offering a flexible scale of fine-grained trust between the admitted users. Note, however, that in this case a user may be assigned with s identities, a user may hold (at most) two points (or identities) on some polynomial $f_{ij}(\cdot, \cdot)$ (one obtained from initiator I_i and the other obtained from I_j). Another possibility is to allow a user to reuse its identity obtained from the first time it is admitted; this way, a user only holds at most one point (or identity) on any polynomial $f_{ij}(\cdot, \cdot)$. This may require a user to authenticate itself to the newly encountered initiator, which can be done because the initiator was also issued credentials by the initiators (including itself). Note that this cost is extra to the cost caused by the enforcement of the policy.

The above scheme adopts a building-block due to Blom [1984], but there is actually a class of instantiations. In what follows we show that the building block can be based on many constructs. First, the schemes of Eschenauer and Gligor [2002], Chan et al. [2003], and Zhu et al. [2003] can be adopted in a plug-and-play fashion, by simply letting each pair of initiators (i, j) share a key pool.

Similarly, the building-block can be respectively based on the probabilistic key predistribution schemes of Du et al. [2003], and Liu and Ning [2003] and the deterministic key predistribution schemes [Camtepe and Yener 2004; Lee and Stinson 2004; Chan and Perrig 2005] in a plug-and-play fashion, by simply letting each pair of initiators (i, j) share a building-block key distribution scheme (e.g., a polynomial pool).

Second, the framework can also be configured to accommodate the online server based scheme [Needham and Schroeder 1978], although such a construction is less interesting because of its high communication cost and the potential to trap into the routing-security interdependence cycle.

5. ON SELECTING PARAMETERS

In the last two sections we explored the security properties of the two newly proposed distributed bootstrapping schemes, one based on public key cryptosystems and the other based on symmetric key cryptosystems. In the latter case, we also investigated how the degree of the polynomials should be determined so as to ensure security in the presence of some powerful adversary. In this section, we explore how the following important parameters should be selected: (1) the time at which the initiators should downgrade their roles to normal users, and (2) the number of initiators.

In order to simplify the analysis, we use discrete time model to approximate continuous time processes on a rectangular grid topology, denoted by $\mathbb{G}(\mathbb{V}, \mathbb{E})$, where $\mathbb{V}$ is the set of vertices with $|\mathbb{V}| = N$ and $\mathbb{E}$ is the set of edges with $|\mathbb{E}| = m$. Recall that $\mathbb{U}$ is the user population and $\mathbb{I}$ is the set of initiators, where $|\mathbb{U}| = n$ and $|\mathbb{I}| = \ell$. Therefore the two parameters can be elaborated as follows:

(1) **The time step at which the initiators should downgrade their roles to normal users**. Given the goal of admitting n outside users, denote by $t_M^{(1\times s;\ell)}$ the convergence time that a user obtains s credentials from s (out of the ℓ) initiators, and by $t_M^{(n\times s;\ell)}$ the convergence time that all the n users obtain s credentials from s (out of the ℓ) initiators, where $1 \le s \le \ell$. Ideally, the initiators should downgrade their roles no earlier than $t_M^{(n\times s;\ell)}$.

(2) **The desired number of initiators**. This is important because an initiator may use a different type of tamper-resistant hardware module. It would be ideal that

$$\ell = \min\left\{\ell' : \left(\forall \ell_1 < \ell' : t_M^{(n\times s;\ell_1)} >> t_M^{(n\times s;\ell')}\right) \wedge \left(\forall \ell_2 > \ell' : t_M^{(n\times s;\ell_2)} \approx t_M^{(n\times s;\ell')}\right)\right\}.$$

Since the parameters would depend on the users' mobility behavior, we consider two mobility models, the Random Walk one and the (Restricted) Random Waypoint one. In both cases, we assume that users start their movement from a stationary distribution; that is, the distances between initial positions of the users are even. Since the former case is simpler, we are able to conduct an analytic analysis. For the latter, we conduct a simulation study to show how the above parameters may be determined. Notice that since the parameters abstracted away implementation details (e.g., whether admission is based on

public key or symmetric key cryptosystems), the results are equally applicable to any concrete instantiations of the framework.

5.1 The Case of Random Walk Mobility Model

In this mobility model, users perform independent random walks on a rectangular toroidal grid. We assume that the grid has continuous boundary conditions, meaning that its boundary vertices are connected to the boundary vertices on the opposite side of the grid. We also assume that a user obtains a credential from the initiator when they are located on the same vertex. Specifically, we represent user movement on the grid as a Markov chain with state space $\mathbb{S}$, where each state which consists of $\mathbb{V}$ in a fixed order (i.e., a state is a vector of length N). Denote by $S_i \in \mathbb{S}$ the state in which vertex $i \in \mathbb{V}$ is occupied. A user initially located at state $S_i \in \mathbb{S}$ moves at each step with an equal probability (i.e., $\frac{1}{4}$) to one of its neighbor states. Then, the position of user u at time t is an independent Markov process with a stationary distribution $\pi = \{\pi_i : S_i \in \mathbb{S}\}$. In the specific case of grid, $\pi_i = \frac{1}{N}$, $\forall S_i \in \mathbb{S}$. Therefore, the transition probabilities of this chain are given by:

$$p_{ij} = \begin{cases} 1/4 & \text{if } (i, j) \in \mathbb{E} \\ 0 & \text{otherwise} , \end{cases}$$

where $p_{ij} = \Pr\left[X_u(t+1) = S_j | X_u(t) = S_i\right]$, $\forall S_i, S_j \in \mathbb{S}$, and $X_u(t)$ is the position of user $u \in \mathbb{U}$ at time t.

Definition 5.1. (average and worst-case mean meeting times) The average mean meeting time $\overline{t_M}^{(1\times1;1)}$ of a user u and an initiator v is the average of the expected meeting times, where average is over all pairs of vertices at which u and v start their walks. That is,

$$\overline{t_M}^{(1\times1;1)} = \frac{1}{N^2} \sum_{i,j\in\mathbb{S}} \mathsf{E}\left[t_M^{(1\times1;1)} | X_u(0) = S_i, X_v(0) = S_j\right]$$

where $t_M^{(1\times1;1)} = \min\{t : X_u(t) = X_v(t)\}$ is the first meeting time of u and v.

The worst case mean meeting time $\tau_M^{(1\times1;1)}$ is defined as the maximum of the expected meeting time of u and v, where maximum is over all pairs of vertices at which u and v start their walks. That is,

$$\tau_M^{(1\times1;1)} = \max_{S_i,S_j\in\mathbb{S}} \mathsf{E}\left[t_M^{(1\times1;1)} | X_u(0) = S_i, X_v(0) = S_j\right]$$

PROPOSITION 5.1. ($\overline{t_M}^{(1\times1;1)}$ and $\tau_M^{(1\times1;1)}$) The average- and wort-case mean meeting times for a user u and an initiator v are:

$$\overline{t_M}^{(1\times1;1)} \approx 0.17N \log N; \qquad \tau_M^{(1\times1;1)} \approx 0.183N \log N$$

PROOF. We make use of the fact that the u's and v's joint Markov chain $(X_u - X_v)(t)$ behaves precisely as $X_u(2t)$, a single random walk with transition time doubled. Hence, the expected meeting time of u and v, given that the nodes start from vertices i and j, respectively, is exactly half of the expected hitting time (t_H) of a single node, starting from i to hit j [Doyle and Snell

2001]. Therefore, $\overline{t_M}^{(1\times1;1)} = \frac{1}{2}\cdot\overline{t_H}$ and $\tau_M^{(1\times1;1)} = \frac{1}{2}\cdot\tau_H$, where $\overline{t_H}$ and τ_H are the average and worst-case mean hitting times, respectively. From [Ellis], we have that $\overline{t_H} \approx 0.34N\log N$ and $\tau^* \approx 0.73N\log N$ as $N \to \infty$, where τ^* is the maximum mean commute time. This approximation is valid already for $N \geq 25$ [Ellis]. Due to the torus symmetry, we have that $\tau_H = \frac{1}{2}\cdot\tau^*$. Thus, $\overline{t_M}^{(1\times1;1)} \approx 0.17N\log N$ and $\tau_M^{(1\times1;1)} \approx 0.183N\log N$. □

PROPOSITION 5.2. ($\overline{t_M}^{(1\times1;\ell)}$ and $\tau_M^{(1\times1;\ell)}$) The average- and worst-case mean times that a user u meets an initiator are:

$$\overline{t_M}^{(1\times1;\ell)} \approx \frac{0.17N\log N}{\ell}; \qquad \tau_M^{(1\times1;\ell)} \approx \frac{0.183N\log N}{\ell}$$

PROOF. We use a very intuitive argument: the Markov chain $\min\{(X_u - X_{v_1})(t), ..., (X_u - X_{v_s})(t)\}$ of u and initiators $v_1, ..., v_s$ runs at an s-times faster rate than $(X_u - X_v)(t)$. Then, the conclusion holds based on Proposition 5.1. □

PROPOSITION 5.3. (Expected meeting time $\overline{t_M}^{(1\times s;\ell)}$ for u to meet $1 \leq s \leq \ell$ initiators)

$$\overline{t_M}^{(1\times s;\ell)} = \overline{t_M}^{(1\times1;1)}\sum_{i=1}^{s}\frac{1}{i}$$

PROOF. The hitting time distribution of an ergodic Markov chain can be approximated by an exponential distribution of the average mean hitting time of the same chain [Aldous and Fill 2000; Shah et al. 2003; Doyle and Snell 2001]. For the meeting time, we can directly apply the same approximation. Thus, the cumulative distribution function (cdf) of the meeting time of a node and an initiator is given by

$$\Pr\left[t_M^{(1\times1;1)} \leq t\right] \approx 1 - e^{\frac{-t}{kN\log N}}$$

where $k = 0.17$. Note that this is a continuous time representation of the discrete Markov chain, but the result equally applies to the discrete time case.

Since the event that user u meets initiator I is independent of the event that it meets any other initiator, the distribution of $t_M^{(1\times s;\ell)}$ is given by

$$\Pr\left[t_M^{(1\times s;\ell)} \leq t\right] \approx \left(1 - e^{\frac{-t}{kN\log N}}\right)^s$$

Denote by $\lambda = \frac{1}{kN\log N}$, then we have

$$\begin{aligned}
\overline{t_M}^{(1\times s;\ell)} &= \mathsf{E}[\max_{I\in\mathbb{I}} t_M(u,I)] = \int_0^\infty \left[1 - \Pr\left[t_M^{(1\times s;\ell)} \leq t\right]\right]dt \\
&= \int_0^\infty \left[1 - \left(1 - e^{-\lambda t}\right)^s\right]dt = \int_0^\infty \left[1 - \sum_{i=0}^{s}\binom{s}{i}(-1)^i e^{-\lambda i t}\right]dt \\
&= \sum_{i=1}^{s}\left[\binom{s}{i}(-1)^{i+1}\frac{1}{\lambda i}\right] = \frac{1}{\lambda}\sum_{i=1}^{s}\frac{1}{i} = \overline{t_M}^{(1\times1;1)}\sum_{i=1}^{s}\frac{1}{i}.
\end{aligned}$$

□

5.1.1 *When should the initiators downgrade their roles to normal users?* In the case of each user only needs to obtain one credential, Proposition 5.2 indicates that the average mean time for any user to obtain at least one credential (i.e., having met at least one initiator) is $\frac{0.17N\log N}{\ell}$. The worst-case mean time for any user to meet at least one initiator is given by $\frac{0.183N\log N}{\ell}$. Therefore, after $\frac{0.183N\log N}{\ell}$ time steps, the initiators can downgrade their role to prevent any attacker from obtaining the management capability. In the case of each user is allowed to obtain multiple credentials, Proposition 5.3 indicates that the worst-case mean time is $0.183N\log N\sum_{i=1}^{s}\frac{1}{i}$. Therefore, after (on average) $0.183N\log N\sum_{i=1}^{s}\frac{1}{i}$ time steps, the initiators can downgrade their roles so as to prevent any attacker from obtaining the management capability.

5.1.2 *How should the number of initiators be determined?* Suppose the network is planned to admit n outside users within a desired time step T. In the case of each user only needs to obtain one credential, Proposition 5.2 indicates that the desired number of initiator is given by $\ell = \frac{0.183N\log N}{T}$. In the case each user is allowed to obtain multiple credentials (this is effectively similar to a threshold cryptosystem-based admission such that a user is fully admitted into the network if it has been admitted, or met, s out of the ℓ initiators), Proposition 5.3 indicates the desired number of initiator is given by $\ell = \frac{0.183N\log N\sum_{i=1}^{s}\frac{1}{i}}{T}$. These results show that the number of initiators depends only on the size of the observed area and on the time within which the associations are to be established. This is an expected result since the motions of nodes are mutually independent and the time required for each node to meet with the initiator will therefore be the same (in average).

5.2 The Case of Random Waypoint Mobility Model

Having explored the case of Random Walk mobility model, we now investigate the Random Waypoint mobility model [Johnson 1994; Camp et al. 2002], which is the most common mobility model for mobile ad hoc networks. More specifically, we consider the extended Restricted Random Waypoint mobility model. In the conventional Random Waypoint model, a user (i.e., node) chooses its destination and its speed towards its destination randomly. After arriving at the destination, the node pauses for a certain period of time, and then chooses its new destination and its speed. In the Restricted Random Waypoint model, users move in the same way as in the Random Waypoint model, except that their choice of destination points is restricted to a number of fixed points on the plane with some probability p. This means that with probability p, a user randomly chooses a point from a finite set of destination points, and with probability $1-p$, it chooses as its destination a random point on the plane. This extended model is closer to reality in the sense that users normally do not randomly choose any point on a plane as their destination, but they rather move to some meeting points (e.g., meeting rooms, lounges, restaurants). In this mobility model, an initiator distributes a credential to a node when they are in the security range of each other. The security range is significantly smaller than

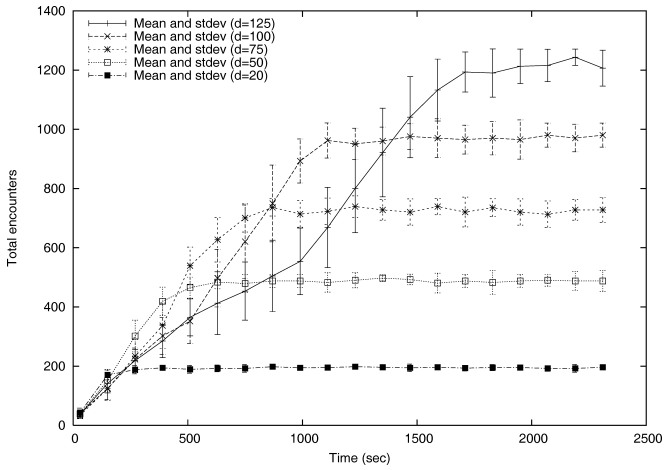

Fig. 3. The total number of encounters with the initiators (i.e., the number of credentials issued) as a function of time and different downgradation conditions (i.e., the number outside users an initiator has encountered). Every point on the graph is the average of 20 runs of the simulation. The vertical bars represent the standard deviations.

the power range of mobile nodes and is the maximum range that is sufficient for the authenticated channel to be set up.

Since we are unable to establish an analytic understanding in this case, we investigate the impact of parameters through simulations. In all simulations, we use the same simulation area, a $100m \times 100m$ square area, we set the number of nodes to $n = 150$ and the number of initiators $\ell = 10$. We observe the distribution of credentials within a bounded rectangle simulation area. The node maximum speed is set to $5m/s$ and the minimum speed to $1m/s$ [Yoon et al. 2003]. The pause time is randomly selected between $0s$ and $120s$. In order to exchange credentials the nodes must be within the security range on an initiator. This security range is set to $10m$. The number of nodes that each initiator will issue credentials to before downgrading its status is denoted by d.

Figure 3 shows the total number of encounters (i.e., the number of credentials issued) with the initiators as a function of time. This is plotted for several different downgradation conditions (i.e., the number of times an initiator will issue credentials to nodes before downgrading its role). When all the initiators have downgraded their roles, no more encounters will happen no matter how much time passes, hence the horizontal lines on the graph. It is also clear from the graph that the number of encounters is more or less linear as a function of time until the initiators begin to downgrade their roles. This means that it doesn't matter if the downgrade is made based on the amount of time that has passed or based on the number of nodes an initiator has encountered.

Figure 4 plots the number of nodes with k or more encounters (i.e., credentials issued) as a function of time. As can be seen on the figure the more credentials each node needs, that is, the higher the value of k, the longer it takes for all the nodes to get the required number of credentials. In this simulation

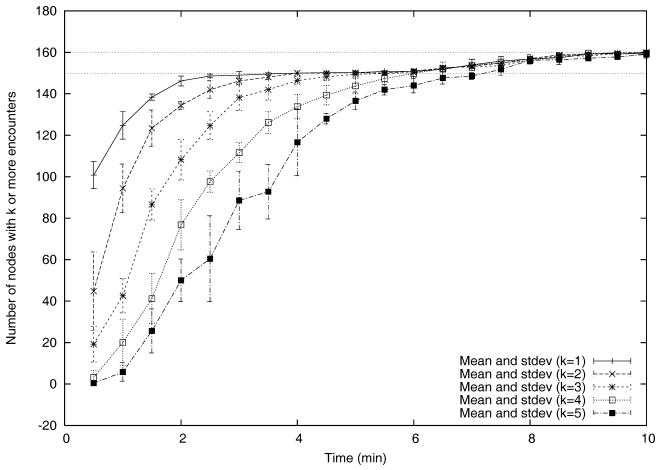

Fig. 4. The number of nodes with k or more credentials as a function of time and different values of k. Every point on the graph is the average of 20 runs of the simulation. The vertical bars represent the standard deviations.

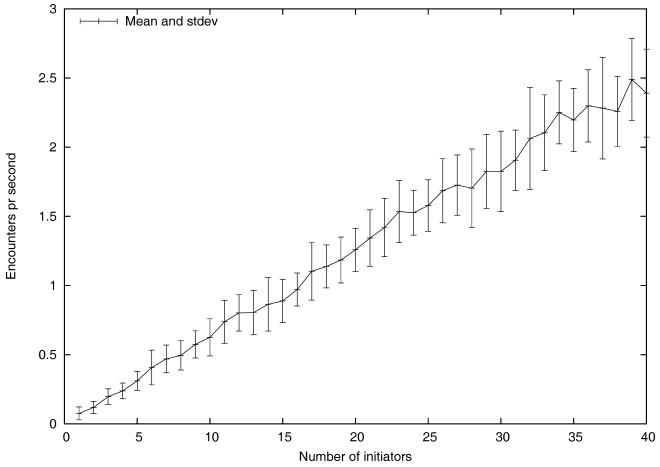

Fig. 5. Encounters per second (i.e., number of credentials issued per second) as a function of the number of initiators. Every point on the graph is the average of 20 runs of the simulation. The vertical bars represent the standard deviations.

the 10 initiators keep their credentials after they downgrade their roles, that is why the number of nodes increases to 160 toward the end.

As one might expect there is a direct connection between the number of initiators and the speed with which encounters are made. Figure 5 plots the number of encounters per second as a function of the number of initiators. We notice that in the Random Walk mobility model we explored in Section 5.1, the impact of more initiators is always linear in terms of meeting times. This is also the case with random Waypoint as shown in Figure 5.

5.2.1 *When should the initiators downgrade their roles to normal users?* We conclude from the above simulations that the number of encounters per normal node is proportional to

$$s = \frac{\ell \cdot d}{n} \quad \Rightarrow \quad d = \frac{s \cdot n}{\ell},$$

which means that the decision about when initiators should downgrade their roles to normal users depends on s: the number of credentials each user needs, n: the number of nodes, d: the number of credentials each initiator issues, and ℓ: the number of initiators.

5.2.2 *How should the number of initiators be determined?* As we discussed above, the number of initiators depends on how quickly we want the credentials exchanged (see Figure 5). The same formula as before can be used to estimate the number of initiators needed.

$$\ell = \frac{s \cdot n}{d}$$

where s is the number of credentials each user needs, n is the number of nodes, and d is the number of credentials each initiator issues before downgrading its role to a normal user.

6. RELATED WORK

From a security perspective, we argued that a democratic threshold cryptosystem based scheme (e.g., Kong et al. [2001], Saxena et al. [2005], and Saxena [2006]) is actually subject to the attack that an adversary may be easily able to compromise a threshold number of users (because they are often geographically close to each other; otherwise the scheme would trap into the routing-security interdependence cycle problem).

Our framework as well as the concrete schemes advocate the use of some hardware modules. Such modules have also been used in some existing schemes including Basagni et al. [2001], and Yi and Kravets [2003]. However, they all are based on the assumption that security of the hardware modules cannot be broken within the lifetime of a MANET. In contrast, we rely on the trust hardware modules only for that they cannot be broken within a much shorter period of time, because the initiators can downgrade their roles by securely erasing the cryptographic materials corresponding to their authority functionalities. Such a prudent engineering practice is important because tamper-resistance is still heuristic [Anderson and Kuhn 1996]. The framework as well as the concrete schemes advocate the use of mobility to help security. Mobility has been exploited to achieve better functionalities [Grossglauser and Tse 2002; Grossglauser and Vetterli 2003] and security [Stajano and Anderson 1999; Asokan and Ginzboorg 2000; Hubaux et al. 2001; Balfanz et al. 2002; Capkun et al. 2003; Capkun et al. 2003; Čagalj et al. 2006].

We analyzed existing bootstrapping schemes ([Zhou and Haas 1999; Kong et al. 2001; Yi and Kravets 2003; Eschenauer and Gligor 2002; Chan et al. 2003; Du et al. 2003; Liu and Ning 2003; Zhu et al. 2003; Camtepe and Yener 2004; Lee and Stinson 2004; Chan and Perrig 2005; Saxena et al. 2005]) that

are closely related to the present paper, with an emphasis on the assumptions upon which they were built. We also showed how the newly established framework can be instantiated to accommodate these schemes.

Finally, we notice that the attempt for distributed key management based on symmetric key cryptosystems made in Chan [2004], which may be seen as a distributed version of Eschenauer and Gligor [2002], was shown to be fundamentally flawed because the flaw is seemingly very difficult to overcome (if not impossible) [Wu and Wei 2005].

7. CONCLUSION AND OPEN PROBLEMS

We presented a framework for distributed and secure bootstrapping of MANETs. The framework is policy neutral, and accommodates existing bootstrapping schemes. Moreover, the framework leads to two (classes of) new schemes that avoid strong assumptions made in existing schemes. The new schemes deploy a small number of trusted initiators that act as the distributed authorities for admitting outside users while exploiting nodes mobility. The initiators are equipped with trusted hardware modules, but the reliance on the tamper-resistance of the modules is minimized. Both schemes are analyzed in terms of the desired properties as well as parameter selections.

This study inspires some interesting open problems. First, how should we extend the bootstrapping framework to accommodate other functions (e.g., detecting and revoking compromised nodes, ensuring secure routing after security associations are established)? A holistic framework is necessary for understanding and adequately addressing MANETs security. Second, we are in lack of a formalism whereby we can rigorously measure or quantify the security of the whole system in the presence multiple compromised users. Traditional cryptographic primitive- or protocol-oriented formalism is not sufficient in this context.

ACKNOWLEDGMENT

We thank the anonymous reviewers for their insightful comments that helped improve the article.

REFERENCES

ALDOUS, D. J. AND FILL, A. 2000. Markov chains on graphs. *Manuscript under preparation*.

ANDERSON, R. AND KUHN, M. 1996. Tamper resistance - a cautionary note. In *Proceedings of The 2nd USENIX Workshop on Electronic Commerce*. The USENIX Association, Oakland, CA, 1–11.

ASOKAN, N. AND GINZBOORG, P. 2000. Key management in ad hoc networks. *Comput. Comm.*, *23*, 1627–1637.

BALFANZ, D., SMETTERS, D., STEWART, P., AND WONG, H. 2002. Talking to strangers: Authentication in ad-hoc wireless networks. In *Proceedings of the Network and Distributed Systems Security Symposium (NDSS'02)*. The Internet Society, San Diego, CA.

BASAGNI, S., HERRIN, K., BRUSCHI, D., AND ROSTI, E. 2001. Secure pebblenets. In *Proceedings of the ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc'01)*. ACM Press, 156–163.

BECHLER, M., HOF, H., KRAFT, D., PAHLKE, F., AND WOLF, L. 2004. A cluster-based security architecture for ad hoc networks. In *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'04)*.

BELLARE, M. AND ROGAWAY, P. 1996. The exact security of digital signatures - how to sign with rsa and rabin. In *Proceedings of the International Conference on the Theory and Application of Crytographic Techniques (EUROCRYPT'96)*. 399–416.

BLOM, R. 1984. An optimal class of symmetric key generation systems. In *Proceedings of the International Conference on the Theory and Application of Crytographic Techniques (EUROCRYPT'84)*. 335–338.

BLUNDO, C., DESANTIS, A., HERZBERG, A., KUTTEN, S., VACCARO, U., AND YUNG, M. 1992. Perfectly-secure key distribution for dynamic conferences. In *Proceedings of the Annual International Cryptology Conference (CRYPTO'92)*, E. F. Brickell, Ed. Springer-Verlag, 471–486. Lecture Notes in Computer Science No. 740.

BRANDS, S. AND CHAUM, D. 1994. Distance-bounding protocols. In *Proceedings of the Workshop on the Theory and Application of Cryptographic Techniques on Advances in Cryptology (EUROCRYPT'94)*. Springer-Verlag New York, Inc., 344–359.

ČAGALJ, M. AND HUBAUX, J. P. 2004. Key agreement over a radio link. Tech. Rep. IC/2004/16, EPFL-DI-ICA. January.

ČAGALJ, M., ČAPKUN, S., AND HUBAUX, J.-P. 2006. Key agreement in peer-to-peer wireless networks. *Proceedings of the IEEE*, to appear.

CAMP, T., BOLENG, J., AND DAVIES, V. 2002. A survey of mobility models for ad hoc network research. *Wireless Communications and Mobile Computing: Special Issue on Mobile Ad Hoc Networking: Research, Trends and Applications*, *2*, 5, 483–502.

CAMTEPE, S. AND YENER, B. 2004. Combinatorial design of key distribution mechanisms for wireless sensor networks. In *Proceedings of the 9th European Symposium on Research Computer Security (ESORICS'04)*. Lecture Notes in Computer Science, vol. 3193. 293–308.

CAPKUN, S., BUTTYAN, L., AND HUBAUX, J. 2003. Self-organized public-key management for mobile ad hoc networks. *IEEE Transactions on Mobile Computing*, *2*, 1, 52–64.

CAPKUN, S., HUBAUX, J., AND BUTTYAN, L. 2003. Mobility helps security in ad hoc networks. In *Proceedings of the 4th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC'03)*. ACM Press, 46–56.

CHAN, A. 2004. Distributed symmetric key management for mobile ad hoc networks. In *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'04)*.

CHAN, H. AND PERRIG, A. 2005. Pike: Peer intermediaries for key establishment in sensor networks. In *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'05)*.

CHAN, H., PERRIG, A., AND SONG, D. 2003. Random key predistribution schemes for sensor networks. In *IEEE Symposium on Security and Privacy 2003 (SP'03)*. IEEE Computer Society, 197–214.

DOUCEUR, J. 2002. The sybil attack. In *Proceedings of the First International Workshop on Peer-to-Peer Systems (IPTPS'01)*. Springer-Verlag, London, UK, 251–260.

DU, W., DENG, J., HAN, Y., AND VARSHNEY, P. 2003. A pairwise key pre-distribution scheme for wireless sensor networks. In *Proceedings of the 10th ACM Conference on Computer and Communications Security (CCS'03)*. ACM, 42–51.

ELLIS, R. Torus Hitting Times Project. http://www.math.tamu.edu/∼rellis/comb/torus/torus.html.

ESCHENAUER, L. AND GLIGOR, V. 2002. A key-management scheme for distributed sensor networks. In *Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS'02)*. 41–47.

GOLDWASSER, S., MICALI, S., AND RIVEST, R. 1988. A digital signature scheme secure against adaptive chosen-message attacks. *SIAM J. Comput.*, *17*, 2 (Apr.), 281–308.

GOODRICH, M., SIRIVIANOS, M., SOLIS, J., TSUDIK, G., AND UZUN, E. 2006. Loud and clear: Human-verifiable authentication based on audio. In *Proceedings of IEEE International Conference on Distributed Computing Systems (ICDCS'06)*.

GROSSGLAUSER, M. AND TSE, D. 2002. Mobility increases the capacity of ad hoc wireless networks. *IEEE/ACM Trans. Netw.*, *10*, 4, 477–486.

GROSSGLAUSER, M. AND VETTERLI, M. 2003. Locating nodes with ease: Mobility diffusion of last encounters in ad hoc networks. In *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'03)*.

HOEPMAN, J. 2004. The ephemeral pairing problem. In *Proceedings of Financial Cryptography (FC'04)*. Lecture Notes in Computer Science, vol. 3110. 212–226.

HU, Y.-C., PERRIG, A., AND JOHNSON, D. B. 2002. Ariadne: a secure on-demand routing protocol for ad hoc networks. In *Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM'02)*. ACM Press, 12–23.

HUBAUX, J., BUTTYAN, L., AND CAPKUN, S. 2001. The quest for security in mobile ad hoc networks. In *Proceedings of the 2nd ACM International Symposium on Mobile Ad Hoc Networking & Computing (MOBIHOC'01)*. ACM Press, 146–155.

KONG, J., ZERFOS, P., LUO, H., LU, S., AND ZHANG, L. 2001. Providing robust and ubiquitous security support for mobile ad hoc networks. In *9th International Conference on Network Protocols (ICNP'01)*. IEEE Computer Society, 251–260.

KRAWCZYK, H. 2003. Sigma: The "sign-and-mac" approach to authenticated Diffie-Hellman and its use in the ike-protocols. In *Proceedings of the Annual International Cryptology Conference (CRYPTO'03)*. Lecture Notes in Computer Science, vol. 2729. 400–425.

LEE, J. AND STINSON, D. 2004. Deterministic key predistribution schemes for distributed sensor networks. In *Proceedings of the 11th International Workshop Selected Areas in Cryptography (SAC'04)*. Lecture Notes in Computer Science, vol. 3357. 294–307.

LIU, D. AND NING, P. 2003. Establishing pairwise keys in distributed sensor networks. In *Proceedings of the 10th ACM Conference on Computer and Communications Security (CCS'03)*. ACM, 52–61.

MCCUNE, J., PERRIG, A., AND REITER, M. 2005. Seeing-is-believing: Using camera phones for human-verifiable authentication. In *IEEE Symposium on Security and Privacy (SP'05)*. 110–124.

NARASIMHA, M., TSUDIK, G., AND YI, J. 2003. On the utility of distributed cryptography in p2p and manets: The case of membership control. In *11th IEEE International Conference on Network Protocols (ICNP'03)*. 336–345.

NEEDHAM, R. M. AND SCHROEDER, M. D. 1978. Using encryption for authentication in large networks of computers. *Comm. ACM*, *21*, 12 (Dec.), 993–999.

OSTROVSKY, R. AND YUNG, M. 1991. How to withstand mobile virus attacks (extended abstract). In *Proceedings of the 10th Annual ACM Symposium on Princiles of Distributed Computing (PODC'91)*. 51–59.

PAPADIMITRATOS, P. AND HAAS, Z. J. 2002. Secure routing for mobile ad hoc networks. In *Proceedings of the Communication Networks and Distributed Systems Modeling and Simulation Conference (CNDS'02)*.

PAPADIMITRATOS, P. AND HAAS, Z. J. 2003. Secure data transmission in mobile ad hoc networks. In *Proceedings of the ACM Workshop on Wireless Security (WiSe'03)*. ACM Press, 41–50.

JOHNSON, D. B. 1994. Routing in Ad Hoc Networks of Mobile Hosts. In *Proceedings of the IEEE Workshop on Mobile Computing Systems and Applications (WMCSA'94)*.

DOYLE, P. G. AND SNELL, J. L. 2001. *Random walks and electric networks*. Number 22. Carus Mathematical Monographs.

SHAH, R. C., ROY, S., JAIN, S., AND BRUNETTE, W. 2003. Data MULEs: Modeling a three-tier architecture for sparse sensor networks. In *Proceedings of the IEEE Workshop on Sensor Network Protocols and Applications (SNPA'03)*.

RIVEST, R., SHAMIR, A., AND ADLEMAN, L. 1978. A method for obtaining digital signatures and public-key cryptosystems. *Comm. ACM*, *21*, 2, 120–126.

SAXENA, N. 2006. Public key cryptography sans certificates in ad hoc networks. In *Proceedings of the 4th International Conference Applied Cryptography and Network Security (ACNS'06)*. Lecture Notes in Computer Science, vol. 3989. 375–389.

SAXENA, N., TSUDIK, G., AND YI, J. 2005. Efficient node admission for short-lived mobile ad hoc networks. In *Proceedings of the 13th IEEE International Conference on Network Protocols (ICNP'05)*. 269–278.

STAJANO, F. AND ANDERSON, R. 1999. The resurrecting duckling: Security issues for ad-hoc wireless networks. In *Proceedings of the 7th International Workshop on Security Protocols*.

TCG. Trusted computing group (tcg) specifications. https://www.trustedcomputinggroup.org/home.

VAUDENAY, S. 2005. Secure communications over insecure channels based on short authenticated strings. In *Advances in Cryptology—Crypto'05*. Springer-Verlag, Berlin, 309–326. Lecture Notes in Computer Science No. 3621.

WU, J. AND WEI, R. 2005. Comments on "distributed symmetric key management for mobile ad hoc networks" from infocom 2004. Cryptology ePrint Archive, Report 2005/008. Available at http://eprint.iacr.org/.

XU, G. AND IFTODE, L. 2004. Locality driven key management architecture for mobile ad hoc networks. In *Proceedings of the IEEE Conference on Mobile Ad Hoc and Sensor Systems (MASS'04)*. 436–446.

YI, S. AND KRAVETS, R. 2003. Moca: Mobile certificate authority for wireless ad hoc networks. In *The 2nd Annual PKI Research Workshop (PKI'03)*.

YOON, J., LIU, M., AND NOBLE, B. 2003. Random Waypoint Considered Harmful. In *Proceedings of the IEEE Conference on Computer Communications (InfoCom'03)*. San Francisco.

ZHOU, L. AND HAAS, Z. 1999. Securing ad hoc networks. *IEEE Networks*, *13*, 6, 24–30.

ZHU, S., XU, S., SETIA, S., AND JAJODIA, S. 2003. Establishing pairwise keys for secure communication in ad hoc networks: A probabilistic approach. In *11th IEEE International Conference on Network Protocols (ICNP'03)*. 326–335.

Received March 2007; revised November 2007; accepted April 2008