Mitigation of Control Channel Jamming under Node Capture Attacks

Patrick Tague, *Student Member, IEEE*, Mingyan Li, *Member, IEEE*, and Radha Poovendran, *Senior Member, IEEE*

Abstract—Availability of service in many wireless networks depends on the ability for network users to establish and maintain communication channels using control messages from base stations and other users. An adversary with knowledge of the underlying communication protocol can mount an efficient denial of service attack by jamming the communication channels used to exchange control messages. The use of spread spectrum techniques can deter an external adversary from such *control channel jamming* attacks. However, malicious colluding insiders or an adversary who captures or compromises system users are not deterred by spread spectrum, as they know the required spreading sequences. For the case of internal adversaries, we propose a framework for control channel access schemes using the random assignment of cryptographic keys to hide the location of control channels. We propose and evaluate metrics to quantify the probabilistic availability of service under control channel jamming by malicious or compromised users and show that the availability of service degrades gracefully as the number of colluding insiders or compromised users increases. We propose an algorithm called GUIDE for the identification of compromised users in the system based on the set of control channels that are jammed. We evaluate the estimation error using the GUIDE algorithm in terms of the false alarm and miss rates in the identification problem. We discuss various design trade-offs between robustness to control channel jamming and resource expenditure.

Index Terms—Wireless multiple access, Control channel jamming, Security, Node capture attacks, Probabilistic metrics.

1 Introduction

Efficient communication in mobile networks requires the use of multiple access protocols allowing mobile users to share the wireless medium by separating user data in any combination of time, frequency, signal space, and physical space. The entire class of multiple access can thus be described by the unifying framework of orthogonal frequency division multiple access (OFDMA) [2]. Allocation of access and resources to mobile users must be periodically updated in order to maintain the efficiency of the multiple access protocol when base station group membership, user demands, and wireless channel conditions are dynamic. Hence, there is a necessary overhead involved in the multiple access protocol to handle the resource allocation to users. This overhead often takes the form of control messages exchanged between mobile users and base stations.

In many systems, dedicated channels are established for the exchange of control messages. These control channels can be used for a wide variety of functions, from topological information propagation for network routing to access control in subscription services. In a cellular system such as GSM [3], [4], [5], for example, base stations and mobile users must coordinate over a variety of control channels in order to perform access control, traffic channel allocation, and inter-cell user handoff. Control channels thus serve as a platform on which higher-level protocol functionality is supported and, hence, as critical points of failure that can be targeted by a malicious adversary in a denial of service (DoS) attack [6].

An adversary with knowledge of the underlying channel access protocol can perform a DoS attack against individual users or local neighborhoods in the mobile network by jamming the communication channels. Moreover, if the access protocol uses a fixed pre-determined schedule for data and control messages, allowing the adversary to distinguish between channels for data and control messages, a *control channel jamming* attack focusing only on the control channels can be mounted with energy savings of several orders of magnitude less than that required to jam all communication channels [7]. The use of jamming-resistant communication protocols such as Direct Sequence Spread Spectrum (DSSS) or Frequency Hopping Spread Spectrum (FHSS) [2], [4] introduce pseudo-randomness into the access schedule by keeping the spreading or hopping sequences, re-

- *P. Tague and R. Poovendran are with the Network Security Lab (NSL), Electrical Engineering Department, University of Washington, Seattle, Washington. Email: {tague,rp3}@u.washington.edu*
- *M. Li is with Boeing Phantom Works and the Network Security Lab (NSL) at the University of Washington. Email: myli@u.washington.edu*

This work is supported in part by the following grants: ONR YIP, N00014-04-1-0479; ARO PECASE, W911NF-05-1-0491; ARL CTA, DAAD19-01-2-001; and ARO MURI, W911NF-07-1-0287. This document was prepared through collaborative participation in the Communications and Networks Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, DAAD19-01-2-0011. The U. S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U. S. Government.
A preliminary version of this material appeared at the 18th Annual IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC'07) [1].
Manuscript received MMM DD, YYYY; revised MMM DD, YYYY.

spectively, unknown to the adversary. It was noted in [8] that the effect of DSSS and FHSS may be further improved by using cryptographic primitives. Alternative anti-jamming techniques include the use of random channel surfing [9] to randomly hop away from jammed channels and re-synchronize on available channels and the use of wormholes [10] to create a channel for reports or alarms from a jammed region.

The above-mentioned anti-jamming techniques consider jamming attacks by an external adversary and are not intended to mitigate jamming by valid network insiders. A set of malicious colluding users or an adversary who captures or subverts network users in a *node capture attack* [11], potentially inserting replicated or fabricated devices into the system [12], is able to bypass the anti-jamming techniques above by assuming the collective roles of the compromised users in the network. For example, a set of malicious colluding users can use the available DSSS or FHSS sequences to perform an efficient jamming attack that follows the corresponding pseudo-random sequence as though it is a fixed schedule. An access protocol which gives the same information to all network users is thus ineffective against DoS attacks by internal adversaries, as a malicious insider has the ability to perform any task of a valid user. Hence, solutions to prevent or mitigate control channel jamming attacks by malicious insiders must make use of the following properties. First, multiple distinct pseudo-random sequences must exist and be held by different users. Second, the set of distinct sequences should exhibit a degree of *cover-freeness* [13] in that at least one of the sequences of each user should be different from the union of the set of sequences held by malicious colluding users with a non-negligible probability to ensure collusion resistance. Finally, the total number of pseudo-random sequences should scale favorably as the number of users increases, suggesting that there exist trade-offs between the cover-free property and the resource efficiency of the protocol.

We thus approach the problem of designing control channel access schemes which allow for efficient reception of control messages while maintaining a degree of independence between the hopping sequences held by different users. In this work, we focus our attention on designing schemes which are robust to control channel jamming attacks by malicious colluding insiders or compromised users.

1.1 Problem Statement and Contributions

In this article, we develop a framework for control channel access schemes that are robust to control channel jamming. Furthermore, we provide techniques for random allocation of control channels to users which yields *graceful performance degradation as the number of compromised users increases*. Our contributions are summarized as follow.

- We develop a correspondence between the problems of key establishment and control channel access in wireless networks and develop a framework for control channel access schemes providing probabilistic availability of control messages using random key assignment.
- We propose metrics of resilience and delay to quantify the probabilistic availability of service and the quality of provided service, respectively, under control channel jamming attacks. We evaluate the proposed metrics by extending existing results for resilience to node capture in wireless networks.
- We propose techniques for the identification and revocation of compromised users by the service provider or a trusted authority that need not be constantly on-line. We formulate the identification problem as a maximum likelihood estimation problem and provide greedy heuristic algorithms using information available to the service provider. We evaluate the identification algorithm by approximating the false alarm and miss rates under the greedy algorithms.
- We provide a simulation study to demonstrate trade-offs that exist between robustness to control channel jamming and resource expenditure which result from the use of random key assignment protocols, serving as a foundation for the design of control channel access schemes.

The remainder of this paper is organized as follows. In Section 2, we state our assumptions about the control channel access model, adversary, and trusted authority. In Section 3, we present a framework for probabilistic control channel access schemes. In Section 4, we propose and evaluate metrics of resilience and delay to characterize the availability of service under control channel jamming. In Section 5, we formulate the identification of compromised users as a statistical estimation problem and analyze the estimation error in terms of the false alarm and miss rates. In Section 6, we provide simulation results and discuss design and implementation trade-offs. Section 7 presents our conclusions and discussion of future work.

2 Model Assumptions

In this section, we state the assumed models for the multiple access protocol and control message structure, adversary, and service provider or trusted authority. We provide a summary of the notation used throughout this work in Table 1.

2.1 Control Message Access Model

We describe the multiple access protocol in terms of the OFDMA framework [2] with separation of signals over orthogonal carrier signals and in time as follows. We let $\Psi = \{\psi_0, \dots, \psi_{M-1}\}$ denote the set of M orthogonal carriers used for wireless communication. We assume that time is slotted and that an initial portion of each time slot is dedicated to control messages. Since we are

TABLE 1
A summary of notation is provided.

Symbol	Definition
$\mathcal{U}, \mathcal{B}$	Set of U users, B base stations
p	Number of time slots in the reuse period
$\mathcal{K}_t$	Set of channel identifiers, or keys, for time slot t
q_t	Number of control channels in time slot t, $\lvert\mathcal{K}_t\rvert$
$\mathcal{K}_{tu}$	Subset of $\mathcal{K}_t$ assigned to user u
m_t	Number of keys per user in time slot t, $\lvert\mathcal{K}_{tu}\rvert$
$\mathcal{C}, c$	Set and number of compromised users, $c = \lvert\mathcal{C}\rvert$
$\mathcal{K}_{t\mathcal{C}}$	Subset of $\mathcal{K}_t$ held by $\mathcal{C}$
$\mathcal{J}_t$	Subset of $\mathcal{K}_{t\mathcal{C}}$ corresponding to jammed channels
θ_t	Probability that each key $k \in \mathcal{K}_{t\mathcal{C}}$ is added to $\mathcal{J}_t$
$r_t(c)$	Slot resilience for time slot t
$r(c)$	Resilience to control channel jamming
$d_t(c)$	Initial-slot delay for time slot t
$d(c)$	Delay due to control channel jamming
$\widehat{\mathcal{C}}$	Estimate of set $\mathcal{C}$ of compromised users
$\mathcal{F}(c)$	False alarm rate in the estimate $\widehat{\mathcal{C}}$ of $\mathcal{C}$
$\mathcal{M}(c)$	Miss rate in the estimate $\widehat{\mathcal{C}}$ of $\mathcal{C}$

focusing on the availability of control messages in this article, we ignore the portion of each time slot dedicated to data. We further partition each time slot t into S sub-slots with duration sufficient to transmit a single control message. Each control channel is thus specified by the time slot t, the sub-slot index $s \in \{0, \ldots, S-1\}$, and the carrier index $j \in \{0, \ldots, M-1\}$ into the set Ψ.

We let $\mathcal{B}$ denote the set of B base stations present in the network. Each base station $b \in \mathcal{B}$ holds the set $\mathcal{K}_t$ of $q_t = |\mathcal{K}_t|$ *control channel identifiers*, each corresponding to a control channel in time slot t. The sub-slot index s and carrier index j corresponding to each identifier $k_t \in \mathcal{K}_t$ are computed using the *control channel locator function* f, assumed to be publicly known. We let $\mathcal{U}$ denote the set of U mobile users in the network and assume that each user $u \in \mathcal{U}$ is within range of at least one base station. Any user $u \in \mathcal{U}$ holding the identifier $k_t \in \mathcal{K}_t$ can locate the corresponding control channel using the function f. We let $\mathcal{K}_{tu}$ denote the subset of $\mathcal{K}_t$ held by user u. We assume that each control message received over the channel identified by k_t carries information relevant to all users in $\mathcal{U}$ holding k_t. This access model is expanded in detail in Section 3.2.

2.2 Adversarial Model

We consider two types of adversaries. First, when malicious insiders collude under the described control channel access structure, they can jam any control channel which can be located using the control channel identifiers they possess. Second, an adversary who captures or subverts system users and assumes their identities in the network can jam control channels using identifiers acquired from compromised users. We let $\mathcal{C} \subseteq \mathcal{U}$ denote the set of *compromised users*, either colluding insiders or those captured by an adversary. For each time slot t, we let $\mathcal{K}_{t\mathcal{C}}$ denote the subset of $\mathcal{K}_t$ collectively held by the compromised users, i.e. $\mathcal{K}_{t\mathcal{C}} = \bigcup_{u \in \mathcal{C}} \mathcal{K}_{tu}$. Furthermore, we let $\mathcal{J}_t$ denote the subset of $\mathcal{K}_{t\mathcal{C}}$ corresponding to control channels in time slot t that are jammed by compromised users.

The case that insiders expose their identifiers to each other and collaboratively choose the subset $\mathcal{J}_t$ is equivalent to that of an adversary in control of multiple compromised users. Alternatively, malicious insiders may independently choose contributions to the overall subset $\mathcal{J}_t$, suggesting that the probability that each key k_t is included in $\mathcal{J}_t$ may increase with the number of compromised users $|\mathcal{C}|$. We let θ_t denote the probability that each identifier $k_t \in \mathcal{K}_{t\mathcal{C}}$ is included in $\mathcal{J}_t$, noting that θ_t may be a function of $|\mathcal{C}|$. In this work, we assume that identifiers in $\mathcal{K}_{t\mathcal{C}}$ are added to $\mathcal{J}_t$ independently with probability θ_t[1].

2.3 Trusted Authority

We assume that a trusted authority[2] (TA) is responsible for the assignment and update of control channel identifiers to users in $\mathcal{U}$ and the identification and revocation of compromised users in the network. We assume that the TA keeps a record of the sets $\mathcal{K}_{tu}$ for each user $u \in \mathcal{U}$ and time slot t and can detect jammed control channels without error, thus recovering the set $\mathcal{J}_t$ for each t. By comparing the collections of sets $\mathcal{K}_{tu}$ and $\mathcal{J}_t$ for various time slots, the TA can determine a set $\widehat{\mathcal{C}}$ of suspected jammers to eliminate from the network. For each time slot t, the set of identifiers $\mathcal{K}_{t\widehat{\mathcal{C}}} \subseteq \mathcal{K}_t$ are removed from $\mathcal{K}_t$ and replaced with fresh identifiers. Any user $u \in \mathcal{U} \setminus \widehat{\mathcal{C}}$ holding an identifier in $\mathcal{K}_{t\widehat{\mathcal{C}}}$ is assigned the corresponding fresh identifiers. We assume that a mechanism for secure key refresh exists and do not further address the key refreshing protocol in this article. We note that, unless the estimation $\widehat{\mathcal{C}}$ of $\mathcal{C}$ is perfect, which is unlikely given the intelligent adversary model, it is possible that valid users in $\mathcal{U} \setminus \mathcal{C}$ will be eliminated from the network or that compromised users in $\mathcal{C}$ who participate in control channel jamming may not appear in $\widehat{\mathcal{C}}$ as suspected jammers.

Our approach does not require constant presence of the TA to oversee the network. In fact, the TA may only be available occasionally to perform the identification and elimination steps by recording jamming evidence $\mathcal{J}_t$, computing the set $\widehat{\mathcal{C}}$ of suspected jammers, and refreshing the control channel identifiers for the remaining users in $\mathcal{U} \setminus \widehat{\mathcal{C}}$. When the adversary compromises system users over an extended duration of time, the random or deterministic *identification interval* between successive identification steps by the TA impacts the total number of compromised users $|\mathcal{C}|$ for a given identification step and the ability for the TA to identify those users with the estimate $\widehat{\mathcal{C}}$.

[1] This assumption is information-theoretically minimal in that the entropy of the set $\mathcal{J}_t$ is maximized for a given θ_t [14].

[2] The TA can be a service provider or a subset of the base stations in $\mathcal{B}$, for example.

3 Random Key Assignment Framework for Control Channel Access

In this section, we develop a correspondence between the problems of control channel access and symmetric key assignment. We show that efficient and robust control channel access can be provided using random key assignment, yielding a framework for probabilistic control channel access schemes.

3.1 Problem Mapping

We provide a one-to-one mapping between control channel access for multiple users in a single time slot and the assignment of symmetric keys to network nodes for use in cryptographic protocols. The mapping is formalized by constructing a bipartite graph [15] which uniquely maps between control channel access schemes and symmetric key assignment schemes.

For a given time slot t, let $G_t = (\mathcal{U} \cup \mathcal{B}, \mathcal{K}_t, E_t)$ be a bipartite graph with left vertex set $\mathcal{U} \cup \mathcal{B}$, right vertex set $\mathcal{K}_t$, and edge set $\mathcal{E}_t \subseteq (\mathcal{U} \cup \mathcal{B}) \times \mathcal{K}_t$. The edge (u, k_t) for $u \in \mathcal{U}$ is in E_t if and only if $k_t \in \mathcal{K}_{tu}$, so u can compute the corresponding channel location using the locator function f. A user $u \in \mathcal{U}$ can receive control messages from a base station $b \in \mathcal{B}$ if and only if $(b, k_t) \in E_t$ and $(u, k_t) \in E_t$. Any compromised user $w \in \mathcal{C}$ that also holds the identifier k_t can compute the channel location using the locator function f, so the channel can be jammed if and only if there is at least one such user $w \in \mathcal{C}$ such that $(w, k_t) \in E_t$.

The bipartite graph G_t constructed above is next used to uniquely construct a symmetric key assignment scheme used to establish secure communication in a wireless network. Let $\mathcal{K}_t$ be a set of symmetric cryptographic keys [16], and let $\mathcal{N} = \mathcal{U} \cup \mathcal{B}$ represent the set of network nodes. For each $k_t \in \mathcal{K}_t$, assign the key k_t to node $n \in \mathcal{N}$ if and only if $(n, k_t) \in E_t$. A pair of nodes $n_1, n_2 \in \mathcal{N}$ can communicate securely if and only if there exists at least one key $k_t \in \mathcal{K}_t$ such that both $(n_1, k_t) \in E_t$ and $(n_2, k_t) \in E_t$. If the key k_t is held by any compromised node $w \in \mathcal{C}$, the communication between n_1 and n_2 is insecure against attacks by the adversary (e.g. eavesdropping on encrypted messages). Hence, the secure link is compromised if and only if there is one such user $w \in \mathcal{C}$ such that $(w, k_t) \in E_t$.

The bipartite graph G_t thus provides a one-to-one correspondence between control channel access schemes and symmetric key assignment schemes[3]. Hence, key assignment solutions that provide secure communication which is robust to node capture attacks can be used to design control channel access schemes which are resilient to control channel jamming attacks by compromised users.

3.2 Random Assignment of Control Channel Keys

Using the mapping in Section 3.1, we make use of the symmetric key assignment model in [17] to provide a framework for probabilistic control channel access using random key assignment. The proposed framework can then be used to design control channel access schemes which are robust to jamming by compromised users. For the remainder of this article, we use the term *control channel key* interchangeably with control channel identifier.

As discussed in Section 1, a control channel access scheme is only robust to control channel jamming by compromised users if a user holds keys that are not held by any compromised user with high probability. Due to fact that any users in $\mathcal{U}$ can be compromised, it is necessary to impose a degree of disparity between the sets $\mathcal{K}_{tu}$ of assigned keys. This necessary disparity is seen by noting that if all sets $\mathcal{K}_{tu}$ are equal, for example when all users share a single global key, a single compromised user can jam all control messages. However, increasing the diversity of keys assigned to different users implies that the total number of keys q_t for each time slot t must increase. Increasing the number of keys q_t in time slot t further implies that the key storage and the number of control messages transmitted by each base station in time slot t increase. Hence, inherent trade-offs exist between the robustness to control channel jamming and the efficiency of the protocol in terms of storage and communication overhead. These trade-offs are discussed in Section 6.

The random assignment of control channel keys to system users is described as follows. For each user $u \in \mathcal{U}$ and time slot t, the subset $\mathcal{K}_{tu}$ of m_t keys[4] is randomly selected from $\mathcal{K}_t$ and assigned to u, independent of other users in $\mathcal{U}$. The choice of random key assignment is motivated by the following observations. First, without prior assumptions on the maximum number of compromised users, choosing the parameters of a deterministic key assignment scheme [7] may not be possible. Second, the imposed structure of deterministic key assignment schemes may allow the adversary to learn information about the assignment of keys to users other than those in $\mathcal{C}$, as shown in [18]. In random key assignment, each user is assigned keys independently, so the adversary cannot learn any information about the assignment of keys to users other than those in $\mathcal{C}$.

To maintain finite key storage for each user and base station and to prevent frequent re-assignment of keys to all users in the network, we adopt the periodic reuse of keys in time slots such that in any time slot t, control channels are located using the keys in the assigned subset $\mathcal{K}_{iu} \subseteq \mathcal{K}_i$ for $i \equiv t \pmod{p}$. The *reuse period* p is

[3] We note that, by assumption, each left node $b \in \mathcal{B}$ is joined to all right nodes $k_t \in \mathcal{K}_t$, though the mapping holds regardless of this assumption, suggesting a natural extension of the problem in which each base station b is assigned a subset of $\mathcal{K}_t$ instead of the entire set. Alternatively, the mapping allows for modeling of the case in which base stations are not present and the users organize in an ad-hoc manner. These extensions are not addressed in this article.

[4] The number of assigned keys per time slot can vary among users as m_{tu}, though we do not address this extension in this work. Analytical results can be obtained using techniques in [18].

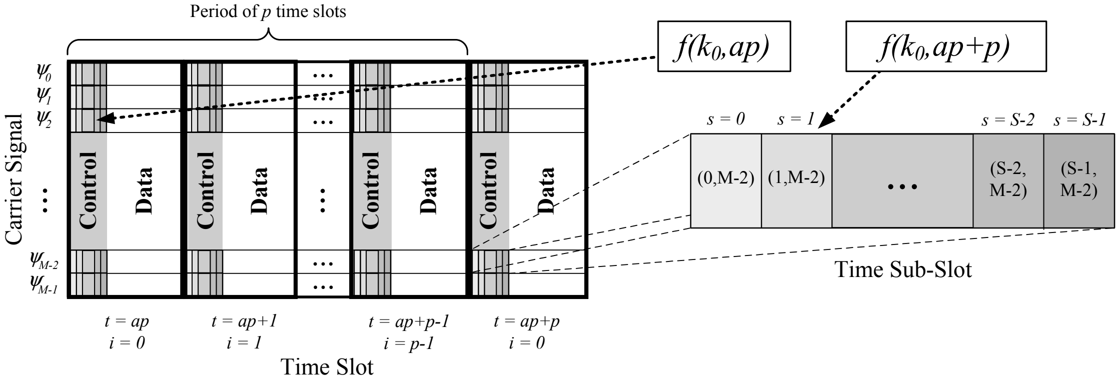


Fig. 1. A control channel access scheme using random key assignment allows for pseudo-random relocation of control channels over time, preventing an adversary from learning via correlation. Each user and base station with a control channel identifier k_i for $i \equiv t \pmod{p}$ locates the corresponding control channel in time slot t as $(s, j) = f(k_i, t)$, where s is a sub-slot index in slot t and j is an index into the set Ψ of carrier signals.

thus a parameter in the design of the control channel access scheme. An additional benefit of the finiteness constraint on the number of distinct time slots is that we can construct the sets $\mathcal{K}_i$ for $i = 0, \ldots, p-1$ to be pairwise disjoint. Furthermore, we assume that the p subsets $\mathcal{K}_{iu}$ for each user u are independently selected, implying that the probabilistic availability of control messages in each time slot is independent.

A consequence of the periodic reuse of keys from each subset $\mathcal{K}_i$ is that control channels will appear at the same location every p time slots if the locator function f depends only on the key k_i. In this case, the adversary may be able to learn the locations of control channels by correlating transmission patterns in corresponding time slots. To prevent transmission correlation, the locator function f must take an additional parameter to vary the control channel location in subsequent periods. In a given time slot t such that $i \equiv t \pmod{p}$, the sub-slot index s and carrier index j are thus given by $(s, j) = f(k_i, t)$. To ensure that distinct control channel keys map to distinct ordered pairs (s, j) with high probability, the locator function f can be implemented using a cryptographic hash function [7], [16]. Fig. 1 illustrates the control channel access scheme.

A control channel access scheme using random key assignment is primarily dependent on the key reuse period p, the number of control channels q_i in each time slot $i = 0, \ldots, p-1$, and the number of control channel keys m_i assigned to each user in $\mathcal{U}$ for use in each time slot $i = 0, \ldots, p-1$. To provide a basis for the design problem, the following sections evaluate the robustness to control channel jamming and the ability to identify and eliminate compromised users from the system.

4 Availability of Control Messages Under Control Channel Jamming

In order to evaluate the effect of control channel jamming by compromised users, we define and evaluate metrics to quantify the probabilistic availability of control messages. We note that users in the proposed control channel access scheme as outlined in Section 3 do not exchange any information about the assigned keys $\mathcal{K}_{iu}$, so the adversary cannot obtain any deterministic information about the key assignment. Intelligent node capture attacks using the techniques proposed in [18] using such information are thus impossible. Hence, the selection of the subset $\mathcal{C} \subseteq \mathcal{U}$ of compromised users is independent of the key assignment, implying that the compromised users are randomly selected by the adversary. We thus define the following metrics to measure the availability of control messages as a function of the number $c = |\mathcal{C}|$ of compromised users, noting that the proposed metrics are computed for the average case.

Definition 1: The *slot resilience* to control channel jamming by $c = |\mathcal{C}|$ compromised users is the probability $r_i(c)$ that a user in $\mathcal{U} \setminus \mathcal{C}$ is able to receive at least one control message in time slot i.

Definition 2: The *resilience* to control channel jamming by $c = |\mathcal{C}|$ compromised users is the probability $r(c)$ that a user in $\mathcal{U} \setminus \mathcal{C}$ is able to receive at least one control message.

Definition 3: The *initial slot delay* due to control channel jamming by $c = |\mathcal{C}|$ compromised users is the average number of time slots $d_i(c)$ that a user in $\mathcal{U} \setminus \mathcal{C}$ must wait to receive a control message when the initial access attempt is made during time slot i. As the delay is infinite for users with no control channel availability, this metric considers only those users able to receive control messages.

Definition 4: The *delay* due to control channel jamming by $c = |\mathcal{C}|$ compromised users is the average number of time slots $d(c)$ that a user in $\mathcal{U} \setminus \mathcal{C}$ must wait to receive a control message, considering only those users able to receive control messages.

In the following sequence of results, we evaluate the resilience and delay metrics using the properties of random control channel key assignment in Section 3.2. We first derive an approximation for the slot resilience

$r_i(c)$. We then prove that the resilience $r(c)$, initial slot delay $d_i(c)$, and delay $d(c)$ can be expressed as a function of the slot resilience $r_i(c)$. The following lemma provides a necessary component in the evaluation of the slot resilience $r_i(c)$.

Lemma 1: When $|\mathcal{C}| = c$, the probability $p_{i,c}(s)$ that $|\mathcal{K}_{iu} \cap \mathcal{J}_i| = s$ for any user $u \in \mathcal{U} \setminus \mathcal{C}$ is approximated as

$$p_{i,c}(s) \approx \binom{m_i}{s} (\theta_i z_{i,c})^s (1 - \theta_i z_{i,c})^{m_i - s}$$

where $z_{i,c} \approx 1 - \left(1 - \frac{m_i}{q_i}\right)^c$.

Proof: Let $p_{c,u}$ denote the probability for a user $u \in \mathcal{U} \setminus \mathcal{C}$ that a particular key $k_i \in \mathcal{K}_{iu}$ is in $\mathcal{J}_i$. Since $\mathcal{J}_i \subseteq \mathcal{K}_{i\mathcal{C}}$, $p_{c,u}$ can be expressed as the product of probabilities $\Pr[k_i \in \mathcal{J}_i | k_i \in \mathcal{K}_{i\mathcal{C}}, k_i \in \mathcal{K}_{iu}]$ and $\Pr[k_i \in \mathcal{K}_{i\mathcal{C}} | k_i \in \mathcal{K}_{iu}]$. The former probability is equal to θ_i by definition. The latter is the probability that at least one of the c compromised users shares the key k_i with user u. Letting $\lambda(k_i)$ denote the number of users in $\mathcal{U}$ holding the key $k_i \in \mathcal{K}$, this probability is approximated by [17, Lemma 6.8] as $1 - \left(\frac{U - \lambda(k_i)}{U-1}\right)^c$, yielding

$$p_{c,u} \approx \theta_i \left(1 - \left(\frac{U - \lambda(k_i)}{U - 1}\right)^c\right). \quad (1)$$

Similar to the result of [17, Theorem 6.9], the average p_c of $p_{c,u}$ over all users $u \in \mathcal{U} \setminus \mathcal{C}$ can be approximated by replacing $\lambda(k_i)$ by its expected value μ_i, yielding $p_c \approx \theta_i z_{i,c}$ where

$$z_{i,c} = 1 - \left(\frac{U - \mu_i}{U - 1}\right)^c. \quad (2)$$

The probability $z_{i,c}$ can be approximated independent of U by noting that $\mu_i = U m_i / q_i$ when keys are assigned randomly, noting that the approximation holds with equality in the limit of large U. Since the keys in $\mathcal{K}_i$ are assigned independently, the probability $p_{i,c}(s)$ satisfies a binomial distribution corresponding to m_i independent trials with success probability p_c. □

The following theorem provides an approximation for the slot resilience $r_i(c)$ using the previous result.

Theorem 1: The slot resilience $r_i(c)$ is approximated as

$$r_i(c) \approx 1 - \theta_i^{m_i} \left(1 - \left(1 - \frac{m_i}{q_i}\right)^c\right)^{m_i}.$$

Proof: By Definition 1, the slot resilience $r_i(c)$ is equal to the probability that $|\mathcal{K}_{iu} \cap \mathcal{J}_i| \neq m_i$ for a user $u \in \mathcal{U} \setminus \mathcal{C}$. Hence, $r_i(c) = 1 - p_{i,c}(m_i)$, where $p_{i,c}(s)$ is given in Lemma 1. □

We next show how the resilience $r(c)$ can be computed as a function of the slot resilience $r_i(c)$ for $i = 0, \ldots, p-1$.

Theorem 2: The resilience $r(c)$ can be computed from the slot resilience $r_i(c)$ for $i = 0, \ldots, p-1$ as

$$r(c) = 1 - \prod_{i=0}^{p-1} (1 - r_i(c)).$$

Proof: A user $u \in \mathcal{U} \setminus \mathcal{C}$ can receive a control message in slot i if and only if $\mathcal{K}_{iu} \nsubseteq \mathcal{J}_i$, an event which occurs with probability $r_i(c)$, by Definition 1. Similarly, a user $u \in \mathcal{U} \setminus \mathcal{C}$ can receive at least one control message in at least one slot if and only if $\mathcal{K}_{iu} \nsubseteq \mathcal{J}_i$ for at least one time slot i, an event which occurs with probability $r(c)$, by Definition 2. The probability $1 - r(c)$ that no channel is available in any slot is given by the product of probabilities $(1 - r_i(c))$ for $i = 0, \ldots, p-1$. Independence of key assignment for different time slots yields the desired result. □

We next show how the initial slot delay $d_i(c)$ and delay $d(c)$ can be expressed as a function of the slot resilience $r_i(c)$. We note that the delay $d(c)$ can be expressed as a weighted sum of the initial slot delays $d_i(c)$ for $i = 0, \ldots, p-1$, where the weight multiplying each $d_i(c)$ is the probability that the initial access attempt is made at time slot i. The probability distribution of delay $d(c)$ can thus be computed as a function of the probability distribution of initial slot delay $d_i(c)$ for $i = 0, \ldots, p-1$. We provide the following result to evaluate the latter probability distribution as a function of the slot resilience $r_i(c)$.

Theorem 3: The probability distribution of $d_i(c)$ is given by

$$\Pr[d_i(c) = \delta] = \gamma_i r_{i+\delta \bmod p}(c) \prod_{d=0}^{\delta-1} (1 - r_{i+d \bmod p}(c))$$

where γ_i is a normalization constant to ensure the summation over $\delta = 0, \ldots, p-1$ equals 1.

Proof: A user must wait δ time slots starting at slot i if and only if there is no control channel available in the first δ time slots beginning at (and including) i and there is a control channel available in slot $(i + \delta \bmod p)$. In each time slot $(i + d \bmod p)$, the probability that no control channel is available is the complement $(1 - r_{i+d \bmod p}(c))$ of the corresponding slot resilience. Independence of key assignment for different time slots yields the desired result. □

In the special case of equal key assignment and jamming parameters for all time slots, the resilience $r(c)$ and the distribution of the delay $d(c)$ can be greatly simplified. The following results illustrate these simplifications.

Theorem 4: When $m_i = m$, $q_i = q$, and $\theta_i = \theta$ for all $i = 0, \ldots, p-1$, the resilience $r(c)$ is approximated as

$$r(c) \approx 1 - \theta^{mp} \left(1 - \left(1 - \frac{m}{q}\right)^c\right)^{mp}.$$

Proof: The result follows directly from Theorem 2 and Theorem 1. □

Theorem 5: When $m_i = m$, $q_i = q$, and $\theta_i = \theta$ for all $i = 0, \ldots, p-1$, the probability distribution of delay $d(c)$ is given by

$$\Pr[d(c) = \delta] = \frac{r_0(c)}{r(c)} (1 - r_0(c))^\delta$$

for $\delta = 0, \ldots, p-1$.

Proof: In the given special case, the slot resilience $r_i(c) = r_0(c)$ for all $i = 0, \ldots, p-1$. Theorem 3 thus yields the probability distribution of $d_i(c)$ as

$$\Pr[d_i(c) = \delta] = \gamma_i r_0(c)(1 - r_0(c))^\delta \tag{3}$$

for $\delta = 0, \ldots, p-1$. The normalization constant $\gamma_i = 1/r(c)$ is obtained by evaluating the finite geometric sum and using the result of Theorem 2. The probability distribution of $d_i(c)$ in (3) is independent of i, so the distribution of $d(c)$ given by any normalized weighted sum of the $d_i(c)$ for $i = 0, \ldots, p-1$ is equal to the distribution of $d_i(c)$. □

To complete the analysis of the delay metric, we compute the expected value $\bar{d}(c)$ of the delay $d(c)$ for the special case approached in Theorem 5. We note that similar techniques can be applied in computing the expected delay in the general case.

Theorem 6: When $m_i = m$, $q_i = q$, and $\theta_i = \theta$ for all $i = 0, \ldots, p-1$, the expected delay $\bar{d}(c)$ is given by

$$\bar{d}(c) = p - 1 + \frac{1}{r_0(c)} - \frac{p}{r(c)}.$$

Proof: The expected value $\bar{d}(c)$ of $d(c)$ is obtained from the result of Theorem 5 using the properties of finite geometric random variables [19]. □

The results obtained in this section can thus be used in the design of control channel key assignment schemes, in particular to balance trade-offs between robustness to control channel jamming and efficiency in terms of key storage and control overhead. These trade-offs in the design process are further discussed in Section 6.

5 Identification of Compromised Users

In this section, we formulate a statistical estimation problem for the identification of compromised users by the TA, constructing a set $\widehat{\mathcal{C}}$ of suspected jammers to eliminate from the network with no knowledge of the number of compromised users $c = |\mathcal{C}|$. Due to the complexity of the resulting identification problem, we propose two algorithms, collectively referred to as GUIDE (Greedy User IDEntification), based on a greedy heuristic which ranks users according to the likelihood of being a compromised user. Finally, we approximate the estimation error resulting from the GUIDE algorithms.

Throughout this section, we denote the parameter vector $(\mathcal{K}_{0u}, \ldots, \mathcal{K}_{(p-1)u})$ as $\mathcal{K}_u$, $(\mathcal{K}_{0\mathcal{C}}, \ldots, \mathcal{K}_{(p-1)\mathcal{C}})$ as $\mathcal{K}_\mathcal{C}$, $(\mathcal{J}_0, \ldots, \mathcal{J}_{p-1})$ as $\mathcal{J}$, and $(\theta_0, \ldots, \theta_{p-1})$ as Θ. Furthermore, we extend the use of logical relations and set cardinalities to parameter vectors in a natural way. For example, we say that $\mathcal{J} \subseteq \mathcal{K}_\mathcal{C}$ if and only if $\mathcal{J}_i \subseteq \mathcal{K}_{i\mathcal{C}}$ for $i = 0, \ldots, p-1$, and we let $|\mathcal{J}| = \sum_{i=0}^{p-1} |\mathcal{J}_i|$.

The information available to the TA and adversary during the attack and identification process are illustrated in Fig. 2. The parameter $\mathcal{K}_u$ is known to the TA for all $u \in \mathcal{U}$ and to the adversary only for $u \in \mathcal{C}$. The evidence $\mathcal{J} \subseteq \mathcal{K}_\mathcal{C}$ is known to the TA and the adversary, but $\mathcal{K}_\mathcal{C}$ is known only to the adversary. The problem of identifying the set $\mathcal{C}$ of jammers is first formulated as a statistical estimation problem in which the TA constructs an estimate $\widehat{\mathcal{C}}$ of $\mathcal{C}$ as a function of the known parameters $\mathcal{J}$ and Θ. When $\Theta = 1$, then $\mathcal{J} = \mathcal{K}_\mathcal{C}$ and the uncertainty in the identification process is greatly reduced. This case was investigated in [1] and is discussed briefly. In addition, as it is quite likely that Θ will not be known to the TA, we vary the heuristic for the case of unknown Θ.

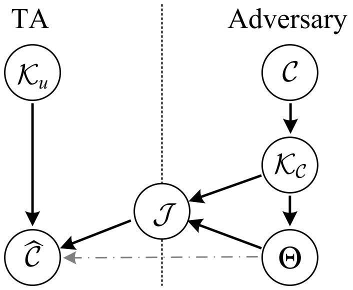


Fig. 2. The information available to the TA and adversary during the attack and identification process is illustrated. The TA has knowledge of the parameters $\mathcal{K}_u$ and $\mathcal{J}$ and uses this available information to construct an estimate $\widehat{\mathcal{C}}$ of $\mathcal{C}$. The adversary has knowledge of the parameters $\mathcal{C}$, $\mathcal{K}_\mathcal{C}$, Θ, and $\mathcal{J}$. The dotted line from Θ to $\widehat{\mathcal{C}}$ indicates that the TA may or may not know Θ.

5.1 Identification with $\Theta = 1$

When $\Theta = 1$, the evidence $\mathcal{J}$ allows the TA to deterministically know $\mathcal{K}_\mathcal{C}$. Hence, the only information the adversary has that the TA does not have is the set $\mathcal{C}$. The TA can thus infer that any user $u \in \mathcal{U}$ holding a key $k_i \notin \mathcal{K}_{i\mathcal{C}}$ for any i cannot be a compromised user. Hence, any user u such that $\mathcal{K}_u \subseteq \mathcal{K}_\mathcal{C}$ is identified as a compromised user, though it is possible that users are falsely identified. This case was addressed in [1] for a similar random key assignment model and in [7] for certain deterministic key assignment schemes.

5.2 Estimation of Compromised User Set $\mathcal{C}$

We formulate the jammer identification problem using statistical estimation by defining the probability $\Pr[\mathcal{C}|\mathcal{J}, \Theta]$ that $\mathcal{C}$ is the set of compromised users responsible for jamming the control channels indicated by the parameters $\mathcal{J}$ and Θ. The estimate which maximizes the probability $\Pr[\mathcal{C}|\mathcal{J}, \Theta]$ is defined as follows.

Definition 5: The *maximum a posteriori (MAP) estimate* [14] $\widehat{\mathcal{C}}$ of the set $\mathcal{C}$ of compromised users is given by

$$\widehat{\mathcal{C}} = \arg\max_{\mathcal{C} \subseteq \mathcal{U}} \Pr[\mathcal{C}|\mathcal{J}, \Theta].$$

An alternate statistical estimation problem can be formulated by defining the likelihood function $\Pr[\mathcal{J}|\mathcal{C}, \Theta]$ that the evidence $\mathcal{J}$ is the outcome of jamming by a given set of compromised users $\mathcal{C}$ and parameter Θ. The estimate which maximizes the likelihood function is defined as follows.

Definition 6: The *maximum likelihood (ML) estimate* [14] $\widehat{\mathcal{C}}$ of the set $\mathcal{C}$ of compromised users is given by

$$\widehat{\mathcal{C}} = \arg\max_{\mathcal{C} \subseteq \mathcal{U}} \Pr[\mathcal{J} | \mathcal{C}, \Theta].$$

The primary difference between the MAP and ML estimates is the availability of prior information about the set $\mathcal{C}$ being estimated, as can be shown using Bayes' Theorem [14]. Since there is no prior information available to the TA about $\mathcal{C}$ and all users are equally likely to be compromised, the MAP and ML estimates are equivalent [14]. The problem of estimating $\widehat{\mathcal{C}}$ can thus be formulated with respect to the likelihood function $\Pr[\mathcal{J}|\mathcal{C}, \Theta]$ characterized by the following results.

Theorem 7: The likelihood function $\Pr[\mathcal{J}|\mathcal{C}, \Theta]$ is given by

$$\Pr[\mathcal{J}|\mathcal{C}, \Theta] = \begin{cases} \prod_{i=0}^{p-1} \theta_i^{|\mathcal{J}_i|} (1-\theta_i)^{|\mathcal{K}_{i\mathcal{C}}| - |\mathcal{J}_i|}, & \text{if } \mathcal{J} \subseteq \mathcal{K}_{\mathcal{C}} \\ 0, & \text{else} \end{cases}.$$

Proof: If $\mathcal{J} \nsubseteq \mathcal{K}_{\mathcal{C}}$, then jamming by compromised users $\mathcal{C}$ could not lead to evidence $\mathcal{J}$. Hence, the likelihood function is non-zero only if $\mathcal{J} \subseteq \mathcal{K}_{\mathcal{C}}$. By assumption in Section 2.2, the sets $\mathcal{J}_i$ for $i = 0, \ldots, p-1$ are selected independently, simplifying the likelihood function as

$$\Pr[\mathcal{J}|\mathcal{C}, \Theta] = \prod_{i=0}^{p-1} \Pr[\mathcal{J}_i|\mathcal{C}, \theta_i]. \tag{4}$$

The dependence of $\mathcal{J}_i$ on $\mathcal{C}$ is in the form of the set $\mathcal{K}_{i\mathcal{C}}$ of keys assigned to compromised users. The likelihood $\Pr[\mathcal{J}_i|\mathcal{C}, \theta_i]$ is thus equal to the probability that independent selection of each element of $\mathcal{K}_{i\mathcal{C}}$ with probability θ_i yields $\mathcal{J}_i$. The likelihood function $\Pr[\mathcal{J}_i|\mathcal{C}, \theta_i]$ is thus given by

$$\Pr[\mathcal{J}_i|\mathcal{C}, \theta_i] = \begin{cases} \theta_i^{|\mathcal{J}_i|} (1-\theta_i)^{|\mathcal{K}_{i\mathcal{C}}| - |\mathcal{J}_i|}, & \text{if } \mathcal{J}_i \subseteq \mathcal{K}_{i\mathcal{C}} \\ 0, & \text{else} \end{cases}. \tag{5}$$

□

In the case that Θ is a constant vector, i.e. $\theta_i = \theta$ for $i = 0, \ldots, p-1$, the likelihood $\Pr[\mathcal{J}|\mathcal{C}, \Theta]$ can be simplified as follows.

Theorem 8: If $\theta_i = \theta$ for $i = 0, \ldots, p-1$, the ML estimate $\widehat{\mathcal{C}}$ of $\mathcal{C}$ is independent of Θ and given by

$$\widehat{\mathcal{C}} = \arg\min_{\substack{\mathcal{C} \subseteq \mathcal{U}, \\ \mathcal{J} \subseteq \mathcal{K}_{\mathcal{C}}}} |\mathcal{K}_{\mathcal{C}}|.$$

Proof: The result of Theorem 7 applied to Definition 6 with $\theta_i = \theta$ yields the estimate

$$\begin{aligned} \widehat{\mathcal{C}} &= \arg\max_{\substack{\mathcal{C} \subseteq \mathcal{U}, \\ \mathcal{J} \subseteq \mathcal{K}_{\mathcal{C}}}} \prod_{i=0}^{p-1} \theta^{|\mathcal{J}_i|} (1-\theta)^{|\mathcal{K}_{i\mathcal{C}}| - |\mathcal{J}_i|} \\ &= \arg\max_{\substack{\mathcal{C} \subseteq \mathcal{U}, \\ \mathcal{J} \subseteq \mathcal{K}_{\mathcal{C}}}} \left(\frac{\theta}{1-\theta} \right)^{|\mathcal{J}|} (1-\theta)^{|\mathcal{K}_{\mathcal{C}}|}. \end{aligned} \tag{6}$$

θ and $\mathcal{J}$ are fixed parameters in the estimation, so the first term in (6) is constant and can be eliminated from the problem. Since $0 \leq (1-\theta) < 1$ and $|\mathcal{K}_{\mathcal{C}}|$ is nonnegative, the maximum is achieved when the exponent is minimized, yielding the desired result independent of Θ. □

```
GUIDE-$\Theta$: Greedy Estimate of $\mathcal{C}$
Given: $\mathcal{J}, \Theta$
$\widehat{\mathcal{C}} \leftarrow \varnothing$
while $\mathcal{J} \nsubseteq \mathcal{K}_{\widehat{\mathcal{C}}}$ do
    $u^* \leftarrow \arg\max_{u \in \mathcal{U} \setminus \widehat{\mathcal{C}}} \Gamma(u|\mathcal{J}, \Theta)$
    $\widehat{\mathcal{C}} \leftarrow \widehat{\mathcal{C}} \cup \{u^*\}$
end while
```

Fig. 3. The algorithm GUIDE-Θ constructs a greedy estimate $\widehat{\mathcal{C}}$ of the set $\mathcal{C}$ of compromised users using the jamming evidence $\mathcal{J}$ and parameter Θ.

We note that, even in the simplified case in Theorem 8, the computation of the ML estimate $\widehat{\mathcal{C}}$ of $\mathcal{C}$ according to Definition 6 requires an exhaustive search through the space of $2^U - 1$ subsets of $\mathcal{U}$, as every subset $\mathcal{C} \subseteq \mathcal{U}$ must be considered in computing the $\arg\max$ and $\arg\min$ functions. Hence, the computation of an estimate $\widehat{\mathcal{C}}$ using maximum likelihood estimation is likely computationally infeasible, unlike maximum likelihood estimation of continuous parameters (e.g. Gaussian noise). We thus shift our attention to the use of heuristics to estimate $\mathcal{C}$.

5.3 Greedy Identification of Jammers - Θ Known

Instead of basing the identification of jammers on the probability $\Pr[\mathcal{C}|\mathcal{J}, \Theta]$ over subsets of $\mathcal{U}$, we base the identification on the probability $\Gamma(u|\mathcal{J}, \Theta)$ that a user $u \in \mathcal{U}$ is a compromised user in $\mathcal{C}$. This heuristic reduces the set estimation problem to a set membership estimation problem. We refer to the identification algorithm as GUIDE, for the **G**reedy **U**ser **IDE**ntification algorithm and first address the case when Θ is known to the TA. In this case, the TA uses a greedy algorithm to construct $\widehat{\mathcal{C}}$ by adding users in decreasing order of probability $\Gamma(u|\mathcal{J}, \Theta)$ until $\widehat{\mathcal{C}}$ satisfies the condition $\mathcal{J} \subseteq \mathcal{K}_{\widehat{\mathcal{C}}}$. This GUIDE-$\Theta$ algorithm is given in Fig. 3. We note that ties can be broken arbitrarily in the $\arg\max$ function, though it is also possible that the $\arg\max$ function can choose an entire subset of users to add to $\widehat{\mathcal{C}}$. This technique and its implications are not addressed in this article.

The probability $\Gamma(u|\mathcal{J}, \Theta)$ for each $u \in \mathcal{U}$ is computed independent of other users in $\mathcal{U}$. In order to compute the desired probability, we define the vector random variable $S_u = (S_{0u}, \ldots, S_{(p-1)u})$ where $S_{iu} = |\mathcal{K}_{iu} \cap \mathcal{J}_i|$ for fixed $\mathcal{K}_{iu}$ and unknown or random $\mathcal{J}_i$. For fixed parameter θ_i, we let $P_{\mathcal{C},i}(s_{iu})$ denote the probability that $S_{iu} = s_{iu}$ given that $u \in \mathcal{C}$ and $P_{\mathcal{U} \setminus \mathcal{C},i}(s_{iu})$ denote the similar probability for $u \in \mathcal{U} \setminus \mathcal{C}$. We further let $P_{\mathcal{C}}(s_u)$ and $P_{\mathcal{U} \setminus \mathcal{C}}(s_u)$ denote the corresponding probabilities that

$S_u = s_u$ for a given vector $s_u = (s_{0u}, \ldots, s_{(p-1)u})$ and fixed Θ.

Lemma 2: The probability $\Gamma(u|\mathcal{J}, \Theta)$ can be expressed as

$$\Gamma(u|\mathcal{J}, \Theta) = \frac{\prod_{i=0}^{p-1} P_{\mathcal{C},i}(s_{iu})}{\prod_{i=0}^{p-1} P_{\mathcal{C},i}(s_{iu}) + \prod_{i=0}^{p-1} P_{\mathcal{U}\setminus\mathcal{C},i}(s_{iu})}.$$

Proof: The keys in $\mathcal{J}_i$ for each time slot i that influence the probability $\Gamma(u|\mathcal{J}, \Theta)$ are only those keys in $\mathcal{K}_{iu} \cap \mathcal{J}_i$ held by u. Since keys are assigned independently and randomly, identification of compromised nodes depends only on the number of keys s_{iu} and not on the specific keys used, so $\Gamma(u|\mathcal{J}, \Theta) = \Gamma(u|s_u, \Theta)$. Using Bayes' Theorem [14] and noting that the event that $u \in \mathcal{C}$ is independent of the parameter Θ, we can express the probability $\Gamma(u|s_u, \Theta)$ as

$$\Gamma(u|s_u, \Theta) = \frac{P_{\mathcal{C}}(s_u) \Pr[u \in \mathcal{C}]}{P_{\mathcal{C}}(s_u) \Pr[u \in \mathcal{C}] + P_{\mathcal{U}\setminus\mathcal{C}}(s_u) \Pr[u \notin \mathcal{C}]}. \quad (7)$$

Since the TA has no prior information about $\mathcal{C}$, every user is equally likely to be compromised, and each possible non-empty set $\mathcal{C}$ is equally likely. This implies that the events $u \in \mathcal{C}$ and $u \notin \mathcal{C}$ are equally likely, so the corresponding factors in (7) cancel. Independence of the key assignment in different time slots implies that the probabilities $P_{\mathcal{C}}(s_u)$ and $P_{\mathcal{U}\setminus\mathcal{C}}(s_u)$ can be factored as

$$P_{\mathcal{C}}(s_u) = \prod_{i=0}^{p-1} P_{\mathcal{C},i}(s_{iu}) \quad (8)$$

$$P_{\mathcal{U}\setminus\mathcal{C}}(s_u) = \prod_{i=0}^{p-1} P_{\mathcal{U}\setminus\mathcal{C},i}(s_{iu}), \quad (9)$$

yielding the desired result. □

To complete the evaluation necessary to perform the GUIDE-Θ algorithm to construct the estimate $\widehat{\mathcal{C}}$, we provide the following lemma to evaluate the probabilities $P_{\mathcal{C},i}(s)$ and $P_{\mathcal{U}\setminus\mathcal{C},i}(s)$.

Lemma 3: The probabilities $P_{\mathcal{C},i}(s)$ and $P_{\mathcal{U}\setminus\mathcal{C},i}(s)$ are given by

$$P_{\mathcal{C},i}(s) = \binom{m_i}{s} \theta_i^s (1-\theta_i)^{m_i - s},$$

$$P_{\mathcal{U}\setminus\mathcal{C},i}(s) = 2^{1-U} \sum_{c=0}^{U-1} \binom{U-1}{c} p_{i,c}(s),$$

where $p_{i,c}(s)$ is approximated by Lemma 1.

Proof: When $u \in \mathcal{C}$ is a compromised user, each key in $\mathcal{K}_{iu}$ appears in the set $\mathcal{J}_i$, and hence in the set $\mathcal{K}_{iu} \cap \mathcal{J}_i$, independently with probability θ_i, yielding

$$P_{\mathcal{C},i}(s) = \binom{m_i}{s} \theta_i^s (1-\theta_i)^{m_i - s}. \quad (10)$$

When $u \notin \mathcal{C}$, the desired probability is the probability that $|\mathcal{K}_{iu} \cap \mathcal{J}_i| = s$ given $u \in \mathcal{U} \setminus \mathcal{C}$. Conditioning this probability on the event that $|\mathcal{C}| = c$ yields the probability $p_{i,c}(s)$ approximated by Lemma 1. Since the TA has no prior information about $\mathcal{C}$, all non-empty subsets of $\mathcal{U} \setminus \{u\}$ are assumed to be equally likely, so $\Pr[|\mathcal{C}| = c | u \notin \mathcal{C}] = 2^{1-U}\binom{U-1}{c}$. □

5.4 Greedy Identification of Jammers - Θ Unknown

When the parameter Θ is unknown to the TA, the GUIDE-Θ algorithm cannot be used, as the probability $\Gamma(u|\mathcal{J}, \Theta)$ cannot be computed. Though it may be possible to construct an estimate $\widehat{\Theta}$ of Θ, we instead suggest replacing the probability $\Gamma(u|\mathcal{J}, \Theta)$ by the alternate selection metric $\kappa_u = \sum_{i=0}^{p-1} s_{iu}$ for each user $u \in \mathcal{U}$. This choice of selection metric is intuitive as users with larger portions of the set of jamming evidence $\mathcal{J}$ should be more likely to appear as compromised users in $\widehat{\mathcal{C}}$. The following result qualifies this replacement as the selection metric.

Theorem 9: The addition of users to $\widehat{\mathcal{C}}$ according to the variables κ_u for $u \in \mathcal{U}$ approximates the addition of users to $\widehat{\mathcal{C}}$ according to the probabilities $\Gamma(u|\mathcal{J}, \Theta)$. Furthermore, if $q_i = q$, $m_i = m$, $\theta_i = \theta$, and c is known, the ordering of $\mathcal{U}$ for the two sets of quantities is identical up to permutation of equal-valued users.

Proof: From Lemma 2 and Lemma 3 with $s_{iu} = |\mathcal{K}_{iu} \cap \mathcal{J}_i|$ and $|\mathcal{C}| = c$ known, $\Gamma(u|\mathcal{J}, \Theta)$ is given by

$$\begin{aligned}\Gamma(u|\mathcal{J}, \Theta) &= \left(1 + \prod_{i=0}^{p-1} z_{i,c}^{s_{iu}} \left(\frac{1-\theta_i z_{i,c}}{1-\theta_i}\right)^{m_i - s_{iu}}\right)^{-1} \\ &= \left(1 + \alpha \prod_{i=0}^{p-1} \beta_i^{s_{iu}}\right)^{-1} \end{aligned} \quad (11)$$

where α and β_i are given by

$$\alpha = \prod_{i=0}^{p-1} \left(\frac{1-\theta_i z_{i,c}}{1-\theta_i}\right)^{m_i}, \quad \beta_i = \frac{1-\theta_i}{z_{i,c}^{-1} - \theta_i}.$$

When $\beta_i = \beta$ for all $i = 0, \ldots, p-1$, as in the special case of $m_i = m$, $q_i = q$, and $\theta_i = \theta$ for all $i = 0, \ldots, p-1$, (11) can be simplified to

$$\Gamma(u|\mathcal{J}, \Theta) = \left(1 + \alpha\beta^{-\kappa_u}\right)^{-1}. \quad (12)$$

The expression in (12) is a monotone increasing function of κ_u, yielding the desired implication. □

The result of Theorem 9 suggests that an alternative to the GUIDE-Θ algorithm in Fig. 3 is given by the GUIDE-κ algorithm in Fig. 4. Moreover, the simplified GUIDE-κ algorithm in Fig. 4 may be a suitable alternative to GUIDE-Θ even if Θ is known, as the required computation is greatly reduced.

5.5 Error in Identification of Compromised Users

In order to evaluate the heuristic estimation problem formulated for the identification of compromised users by the TA, we provide the following metrics of estimation error.

```
GUIDE-$\kappa$: Greedy Estimate of $\mathcal{C}$
Given: $\mathcal{J}$
$\widehat{\mathcal{C}} \leftarrow \varnothing$
while $\mathcal{J} \nsubseteq \mathcal{K}_{\widehat{\mathcal{C}}}$ do
    $u^* \leftarrow \arg\max_{u \in \mathcal{U} \setminus \widehat{\mathcal{C}}} \kappa_u$
    $\widehat{\mathcal{C}} \leftarrow \widehat{\mathcal{C}} \cup \{u^*\}$
end while
```

Fig. 4. The algorithm GUIDE-κ constructs a greedy estimate $\widehat{\mathcal{C}}$ of the set $\mathcal{C}$ of compromised users using the jamming evidence $\mathcal{J}$ and can be used when Θ is unknown to the TA.

Definition 7: The *false alarm rate* $\mathcal{F}(c)$ is the average fraction of jamming suspects in $\widehat{\mathcal{C}}$ which are not compromised users in $\mathcal{C}$ when $|\mathcal{C}| = c$.

Definition 8: The *miss rate* $\mathcal{M}(c)$ is the average fraction of compromised users in $\mathcal{C}$ which do not appear as jamming suspects in $\widehat{\mathcal{C}}$ when $|\mathcal{C}| = c$.

The false alarm and miss rates in Definitions 7 and 8 are approximated using the following sequence of results. For added clarity, the estimation error is approximated with respect to the GUIDE-κ algorithm in Fig. 4 using the quantities κ_u instead of the GUIDE-Θ algorithm in Fig. 3 using the probabilities $\Gamma(u|\mathcal{J}, \Theta)$. We partition the set of random variables κ_u to distinguish between the compromised users in $\mathcal{C}$ and the remaining users in $\mathcal{U} \backslash \mathcal{C}$. The n^{th} largest κ_u values of users in $\mathcal{C}, \mathcal{U} \backslash \mathcal{C}$, and $\mathcal{U}$ are respectively denoted $\kappa_{\mathcal{C}}^{(n)}$, $\kappa_{\mathcal{U} \backslash \mathcal{C}}^{(n)}$, and $\kappa_{\mathcal{U}}^{(n)}$. For clarity, let $\kappa_{\mathcal{U} \backslash \mathcal{C}}^{(n)} = \kappa_{\mathcal{U}}^{(n)} = \kappa_{\mathcal{C}}^{(n)} = \infty$ for $n = 0$, $\kappa_{\mathcal{C}}^{(n)} = 0$ for $n > c$, $\kappa_{\mathcal{U} \backslash \mathcal{C}}^{(n)} = 0$ for $n > U - c$, and $\kappa_{\mathcal{U}}^{(n)} = 0$ for $n > U$. The following lemma characterizes the distributions of the random variables κ_u and $\kappa^{(n)}$, and a proof is provided in the Appendix.

Lemma 4: For $A \in \{\mathcal{C}, \mathcal{U}, \mathcal{U} \setminus \mathcal{C}\}$, the probability $\Phi_A^{(n)}(\kappa|c) = \Pr\left[\kappa_A^{(n)} \geq \kappa \;\middle|\; |\mathcal{C}| = c\right]$ is computed as

$$\Phi_A^{(n)}(\kappa|c) = \sum_{j=n}^{|A|} \binom{|A|}{j} \Phi_A(\kappa|c)^j \left(1 - \Phi_A(\kappa|c)\right)^{|A|-j},$$

from the probabilities $\Phi_A(\kappa|c) = \Pr\left[\kappa_A \geq \kappa \mid |\mathcal{C}| = c\right]$ given by

$$\Phi_{\mathcal{C}}(\kappa|c) = \sum_{k \geq \kappa} \left(P_{\mathcal{C},0} \ast \cdots \ast P_{\mathcal{C},p-1}\right)(k)$$

$$\Phi_{\mathcal{U} \backslash \mathcal{C}}(\kappa|c) = \sum_{k \geq \kappa} \left(p_{0,c} \ast \cdots \ast p_{p-1,c}\right)(k)$$

$$\Phi_{\mathcal{U}}(\kappa|c) = \frac{c}{U} \Phi_{\mathcal{C}}(\kappa|c) + \frac{U-c}{U} \Phi_{\mathcal{U} \backslash \mathcal{C}}(\kappa|c)$$

where $p_{i,c}$ is the probability distribution given by Lemma 1, $P_{\mathcal{C},i}$ is the probability distribution given by Lemma 3, and $\ast$ is the convolution operator for discrete probability distributions.

We note that when $|\widehat{\mathcal{C}}| = \hat{c}$ users appear as jamming suspects with $|\mathcal{C}| = c$ compromised users, the number of falsely accused users $F = |\widehat{\mathcal{C}} \setminus \mathcal{C}|$ and missed compromised users $M = |\mathcal{C} \setminus \widehat{\mathcal{C}}|$ satisfy $\hat{c} = c - M + F$. Hence, the false alarm and miss rates for fixed c are approximated by estimating the distribution of $\hat{c}$ given c and the distribution of F given $\hat{c}$ and c. The probability $p(\hat{c}|c) = \Pr\left[|\widehat{\mathcal{C}}| = \hat{c} \;\middle|\; |\mathcal{C}| = c\right]$ is first estimated using the following result, a proof of which can be found in the Appendix.

Lemma 5: The probability distribution $p(\hat{c}|c)$ of $|\widehat{\mathcal{C}}|$ given $|\mathcal{C}|$ is approximated as

$$p(\hat{c}|c) \approx \sum_{J} P_{\mathcal{J}}(J, c) \frac{Q_{J,c}(J, \hat{c}) - Q_{J,c}(J, \hat{c}-1)}{1 - Q_{J,c}(J, \hat{c}-1)}$$

where $Q_{J,c}(L, \hat{c})$ and $P_{\mathcal{J}}(J, c)$ are defined recursively as

$$\begin{aligned}
Q_{J,c}(L, \hat{c}) &= \sum_{\kappa} \left(\Phi_{\mathcal{U}}^{(\hat{c})}(\kappa|c) - \Phi_{\mathcal{U}}^{(\hat{c})}(\kappa+1|c)\right) \\
&\quad \times \sum_{n=0}^{\kappa} \nu_n Q_{J,c}(L-n, \hat{c}-1), \\
\nu_n &= \binom{\kappa}{n} \left(1 - \frac{L-n}{J}\right)^n \left(\frac{L-n}{J}\right)^{\kappa-n}, \\
P_{\mathcal{J}}(J, c) &= \left(P_{\mathcal{J}}^0(\cdot, c) \ast \cdots \ast P_{\mathcal{J}}^{p-1}(\cdot, c)\right)(J), \\
P_{\mathcal{J}}^i(J_i, c) &= \sum_{k=J_i}^{q_i} \binom{k}{J_i} \theta_i^{J_i} (1-\theta_i)^{k-J_i} P_{\mathcal{K}}^i(k, c), \\
P_{\mathcal{K}}^i(k, c) &= \sum_{n=0}^{m_i} \tau_n P_{\mathcal{K}}^i(k-n, c-1), \\
\tau_n &= \binom{m_i}{n} \left(1 - \frac{k-n}{q_i}\right)^n \left(\frac{k-n}{q_i}\right)^{m_i-n},
\end{aligned}$$

where $\ast$ is the convolution operator for discrete probability distributions.

The probability distribution $p(F|\hat{c}, c) = \Pr\left[|\widehat{\mathcal{C}} \setminus \mathcal{C}| = F \;\middle|\; |\widehat{\mathcal{C}}| = \hat{c}, |\mathcal{C}| = c\right]$, characterizing the behavior of both F and M, is estimated using the following result, a proof of which can be found in the Appendix.

Lemma 6: The probability distribution $p(F|\hat{c}, c)$ of the number of falsely accused users given $|\widehat{\mathcal{C}}|$ and $|\mathcal{C}|$ is approximated as

$$\begin{aligned}
p(F|\hat{c}, c) \approx \sum_{\kappa_1, \kappa_2} &\min\left(1, \frac{\Phi_{\mathcal{C}}^{(\hat{c}-F)}(\kappa_1|c)}{\Phi_{\mathcal{C}}^{(\hat{c}-F)}(\kappa_2|c)}\right) \\
&\times \min\left(1, \frac{\Phi_{\mathcal{U} \backslash \mathcal{C}}^{(F)}(\kappa_2|c)}{\Phi_{\mathcal{U} \backslash \mathcal{C}}^{(F)}(\kappa_1|c)}\right) \\
&\times \left(\Phi_{\mathcal{C}}^{(\hat{c}-F+1)}(\kappa_2|c) - \Phi_{\mathcal{C}}^{(\hat{c}-F+1)}(\kappa_2+1|c)\right) \\
&\times \left(\Phi_{\mathcal{U} \backslash \mathcal{C}}^{(F+1)}(\kappa_1|c) - \Phi_{\mathcal{U} \backslash \mathcal{C}}^{(F+1)}(\kappa_1+1|c)\right).
\end{aligned}$$

Theorem 10: The false alarm rate $\mathcal{F}(c)$ is given by

$$\mathcal{F}(c) = \sum_{\hat{c}=1}^{U} \frac{p(\hat{c}|c)}{\hat{c}} \sum_{F=0}^{\hat{c}} F p(F|\hat{c}, c),$$

where $p(\hat{c}|c)$ is approximated by Lemma 5 and $p(F|\hat{c}, c)$ is approximated by Lemma 6.

Proof: Letting $\mathcal{E}[x]$ denote the expected value of x, Definition 7 suggests that

$$\begin{aligned}\mathcal{F}(c) =& \mathcal{E}\left[\frac{|\widehat{\mathcal{C}} \setminus \mathcal{C}|}{|\widehat{\mathcal{C}}|} \;\middle|\; |\mathcal{C}| = c\right] && (13)\\ =& \sum_{\hat{c}=1}^{U} p(\hat{c}|c)\mathcal{E}\left[\frac{|\widehat{\mathcal{C}} \setminus \mathcal{C}|}{|\widehat{\mathcal{C}}|} \;\middle|\; |\widehat{\mathcal{C}}| = \hat{c}, |\mathcal{C}| = c\right] && (14)\\ =& \sum_{\hat{c}=1}^{U} \frac{p(\hat{c}|c)}{\hat{c}} \sum_{F=0}^{\hat{c}} Fp(F|\hat{c}, c). && (15)\end{aligned}$$

□

Theorem 11: The miss rate $\mathcal{M}(c)$ is given by

$$\mathcal{M}(c) = \frac{1}{c}\sum_{\hat{c}=1}^{U} p(\hat{c}|c) \sum_{F=0}^{\hat{c}} (c + F - \hat{c})p(F|\hat{c}, c),$$

where $p(\hat{c}|c)$ is approximated by Lemma 5 and $p(F|\hat{c}, c)$ is approximated by Lemma 6.

Proof: This result follows from Theorem 10 and the relationship $\hat{c} = c - M + F$. □

6 Numerical Illustration and Design

In this section, we provide simulation results to illustrate design trade-offs, providing a basis for parameter selection in design of the system. We evaluate the metrics derived in Section 4 and 5 and discuss the effect of varying individual design parameters. We simulate the long-term performance of the system as a function of the identification interval of the TA as defined in Section 2.3.

6.1 Simulation Setup

We simulate a network of $U = 250$ users with varying parameter values of p, m_i, and q_i with the jamming probability $\theta_i = 0.9$. For each set of parameters p, m_i, and q_i, we randomly assign p sets of m_i control channel keys to each user from the p sets of q_i keys. For each value of c, the subset $\mathcal{C}$ is randomly selected from the set of users $\mathcal{U}$, and the subsets $\mathcal{J}_i$ of keys used for jamming are selected randomly using the parameter θ_i. For each subset $\mathcal{C}$ of size c, the resilience $r(c)$ is computed as the fraction of the $|\mathcal{U} \setminus \mathcal{C}|$ remaining users that can access at least one control channel. Similarly, the average delay $\bar{d}(c)$, false alarm rate $\mathcal{F}(c)$, and miss rate $\mathcal{M}(c)$ are computed using the GUIDE-Θ algorithm based on the assigned keys, jammed control channels, and compromised users. Each data point in our simulation reflects an average over 100 simulated network and random key assignment instances. The results of the simulation study are illustrated in Figure 5 for four parameter sets. The solid and dashed lines in each plot represent the analytical results derived in Sections 4 and 5, and the symbol-marked points represent the results of the simulation study. As can be seen from Figure 5, the analytical results for the resilience $r(c)$ and the average delay $\bar{d}(c)$ coincide. While the analytical and simulation results for the false alarm rate $\mathcal{F}(c)$ and the miss rate $\mathcal{M}(c)$ disagree at individual values of c, the analytical results provide a reasonable approximation of the error behavior that can be expected.

6.2 Trade-offs in Key Assignment Parameters

We next identify and discuss design trade-offs in key assignment parameters by investigating the impact of individual parameters using the proposed evaluation metrics. We compare resource trade-offs with respect to the required key storage $\sum_{i=0}^{p-1} m_i$ for users in $\mathcal{U}$ and $\sum_{i=0}^{p-1} q_i$ for base stations in $\mathcal{B}$. We note that since each key corresponds to a unique control channel, the communication overhead for base stations is proportional to the base station key storage.

6.2.1 *Time Slots per Reuse Period* p

We first investigate the impact of varying the reuse period p. Intuitively, increasing the number of assigned key sets p increases the disparity of assigned keys between valid and compromised users, increasing the overall robustness to attacks. In terms of resilience, illustrated in Fig. 5(a), the results of Theorems 1 and 4 validate this intuition, as the complement $(1 - r(c))$ of the resilience is a product of p probability terms. However, we note that key storage at each user and base station increases linearly with p. In addition, the average delay increases linearly with p, as seen in Theorem 6 and Fig. 5(b). As seen in Fig. 5(c) and 5(d), increasing p slightly improves the false alarm and miss rates. The effect of varying p is illustrated by comparing the results in Fig. 5 for $p = 4$, $m_i = 4$, and $q_i = 20$ to those of $p = 8$, $m_i = 4$, and $q_i = 20$.

6.2.2 *Control Channels per Time Slot* q_i

We next investigate the impact of varying the number of control channels q_i in each time slot i. Increasing the number of channels implies that users share fewer keys on average, leading to an improvement in robustness to attacks. The results of Theorems 1 and 4 illustrate an inverse dependence on each parameter q_i, suggesting that resilience and delay improve as q_i increases, as seen in Fig. 5(a) and 5(b). Similarly, the identification capabilities of the TA improve with increasing q_i for small c because users are less likely to share keys, as seen in Fig. 5(c) and 5(d). However, key storage at each base station increases linearly with q_i. The effect of varying q_i is illustrated by comparing the results in Fig. 5 for $p = 4$, $m_i = 4$, and $q_i = 20$ to those of $p = 4$, $m_i = 4$, and $q_i = 40$.

6.2.3 *Control Channel Keys per User per Time Slot* m_i

We next investigate the impact of varying the number of control channel keys m_i assigned to each user in time slot i. Increasing m_i implies that users share more keys on average. However, the parameter m_i also appears

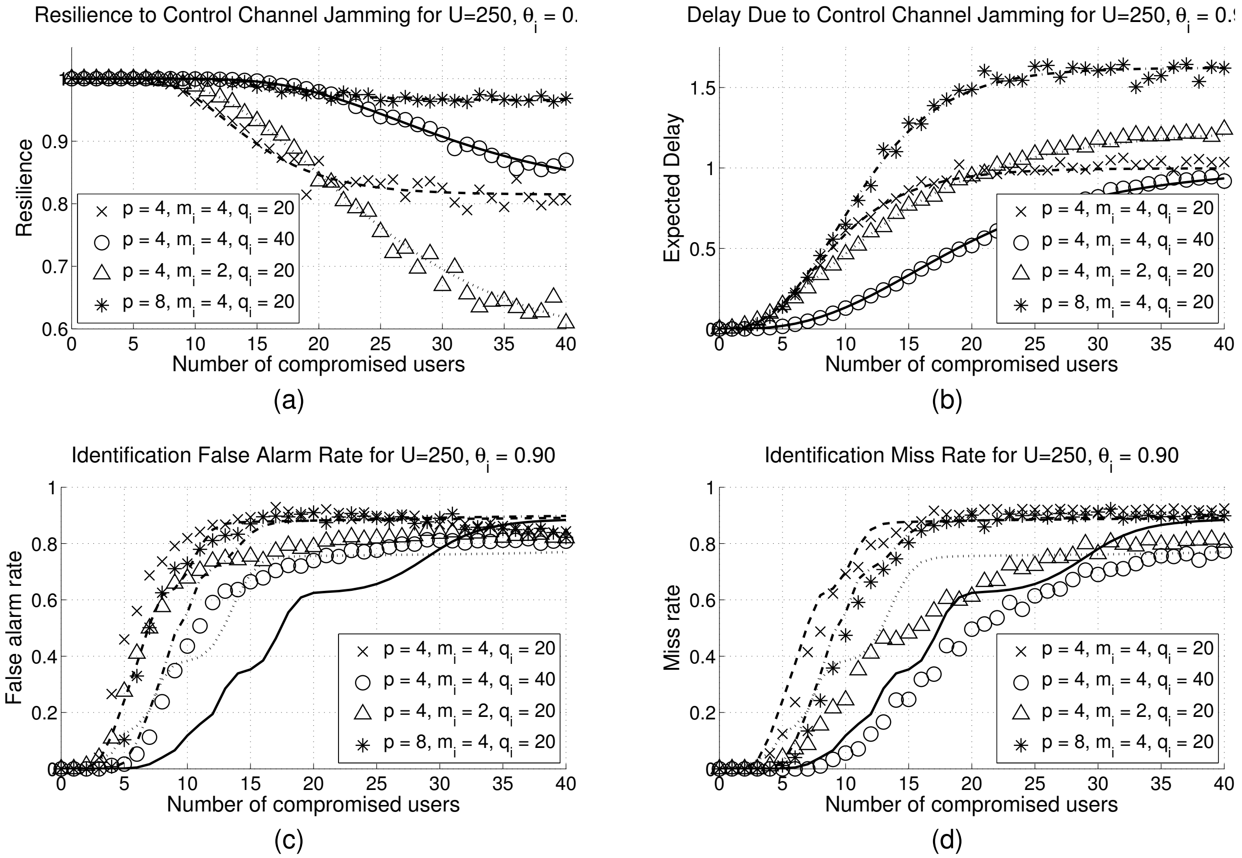


Fig. 5. Variations in the (a) resilience $r(c)$, (b) expected delay $\bar{d}(c)$, (c) false alarm rate $\mathcal{F}(c)$, and (d) miss rate $\mathcal{M}(c)$ are illustrated for a network of $U = 250$ users with varying parameter values of p, m_i, and q_i and a jamming parameter of $\theta_i = 0.9$ using GUIDE-Θ. Solid and dashed lines represent analytical results derived in Sections 4 and 5, and symbol-marked points represent the simulated results averaged over 100 simulated network instances.

as an exponential term in the slot resilience given by Theorem 1. Hence, the effect of varying m_i depends on the values of the parameters p and q_i, suggesting that there exist various trade-offs in varying m_i. The effect of varying m_i is illustrated by comparing the results in Fig. 5 for $p = 4$, $m_i = 2$, and $q_i = 20$ to those of $p = 4$, $m_i = 4$, and $q_i = 20$.

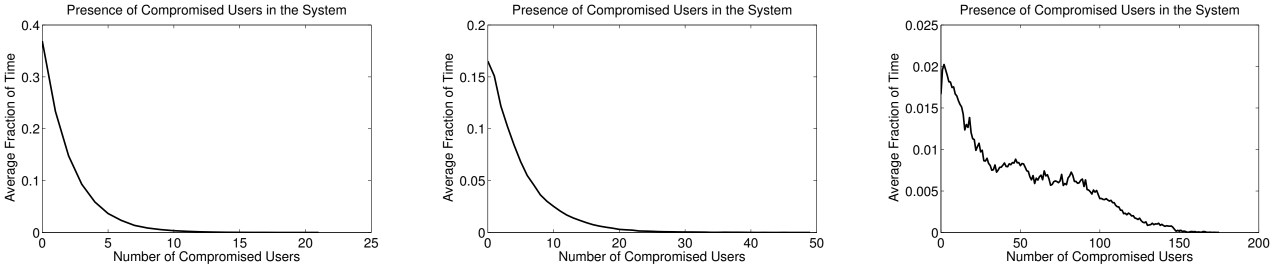
6.2.4 Control Channels per Time Slot q_i and per User m_i

We note that a constant increase in both m_i and q_i allows the designer to take advantage of the increased exponent m_i in the slot resilience given by Theorem 1 without increasing the ratio m_i/q_i. Hence, increasing key storage at users and base stations by a constant leads to an improvement in resilience. The effect of jointly varying q_i and m_i is illustrated by comparing the results in Fig. 5 for $p = 4$, $m_i = 2$, and $q_i = 20$ to those of $p = 4$, $m_i = 4$, and $q_i = 40$.

6.3 Trade-offs in Identification Interval of the TA

As illustrated in Figs. 5(c) and 5(d), the false alarm rate $\mathcal{F}(c)$ and miss rate $\mathcal{M}(c)$ are increasing functions of c when c is small. The performance of the identification process is thus improved if the TA can identify and eliminate the compromised users in the network when there are relatively few of them. Hence, by reducing the length of the identification interval and sampling the system more frequently, the TA can achieve decreased false alarm and miss rates. However, this performance gain comes with a necessary increase in computation and communication overhead for the TA due to the increased rate of collection of jamming evidence $\mathcal{J}$, execution of the GUIDE algorithm, and update of fresh control channel keys to remaining users.

Fig. 6 illustrates the effect of the identification interval on the time-varying number of compromised users in the system using the following simulation setup. A new user is added to the network whenever a user is revoked, so the total number of users U remains constant. After each key reuse period of p time slots, a user in $\mathcal{U} \setminus \mathcal{C}$ is randomly selected and added to $\mathcal{C}$ with probability 0.4. The identification rate determines the frequency with which the TA collects jamming evidence $\mathcal{J}$ and executes the GUIDE algorithm. Fig. 6 plots the normalized histogram of the number of compromised users after each key reuse period to illustrate the long-term average of



(a) (b) (c)

Fig. 6. Identification of compromised users is simulated using the GUIDE-κ algorithm with an adversary compromising users over an extended duration. Each figure plots the normalized histogram of the fraction of 1000 key reuse periods with a given number of compromised users present in the network. The system parameters are chosen as $p = 4$, $m_i = 4$, and $q_i = 20$, and the jamming parameter is chosen as $\theta = 0.9$. The average identification interval is chosen as (a) 5, (b), 10, and (c) 20 key reuse periods.

the number of compromised users in the system. The average identification interval, equal to the inverse of the identification rate, is chosen as 5 periods in Fig. 6(a), 10 periods in Fig. 6(b), and 20 periods in Fig. 6(c). As illustrated, the number of compromised users tends to increase as the identification interval increases, eventually leading to cascading system failure where a majority of the users in the network are compromised.

7 Conclusion

In this article, we addressed the mitigation of control channel jamming by malicious colluding insiders and compromised system users as well as the identification of compromised users without prior knowledge of the number of compromised users in the system. We mapped the problem of control channel access that is robust to jamming by compromised users to the problem of secure key establishment under node capture attacks. Based on the mapping, we proposed a framework for control channel access schemes using random key assignment. We proposed and evaluated metrics for resilience and delay which quantify the availability of control messages under control channel jamming attacks and demonstrated that the use of random key assignment provides graceful degradation in availability as the number of compromised users increases. We formulated the identification of compromised users in the system as a maximum likelihood estimation problem and proposed the GUIDE algorithms using greedy heuristics for jammer identification. We provided an analytical approximation to evaluate the false alarm and miss rates in the identification of compromised users resulting from the GUIDE algorithms. We discussed design trade-offs in the key assignment parameters and the identification interval used by the TA. In future work, we will investigate modifications to the adversary's jamming strategy and the effect on the availability of control messages and the ability to identify compromised users.

Appendix

Proof of Lemma 4: For each $A \in \{\mathcal{C},\mathcal{U},\mathcal{U} \setminus \mathcal{C}\}$, the derivation of $\Phi_A^{(n)}(\kappa|c)$ from $\Phi_A(\kappa|c)$ follows directly using order statistics [20] by noticing that at least n of the $|A|$ independent events are successful, and the success of each event occurs with probability $\Phi_A(\kappa|c)$. When $u \in \mathcal{C}$, the probability $\Phi_{\mathcal{C}}(\kappa|c)$ is the summation over $k \geq \kappa$ of the probability that $\kappa_u = k$. Since $\kappa_u = \sum_{i=0}^{p-1} s_{iu}$, the probability that $\kappa_u = k$ is given by the p-fold convolution evaluated at k of the probability distributions $P_{\mathcal{C},i}$ given by Lemma 3. When $u \in \mathcal{U} \setminus \mathcal{C}$, the corresponding probability is similarly given by the p-fold convolution evaluated at k of probability distributions $p_{i,c}$ given by Lemma 1. The probability $\Phi_{\mathcal{U}}(\kappa|c)$ is computed using the law of total probability and the fact that $\Pr[u \in \mathcal{C}] = c/U$ and $\Pr[u \in \mathcal{U} \setminus \mathcal{C}] = (U - c)/U$. □

Proof of Lemma 5: The probability $p(\hat{c}|c)$ is computed as the probability that the greedy algorithm stops after adding $\hat{c}$ users to $\widehat{\mathcal{C}}$ when $|\mathcal{C}| = c$. This is thus equivalent to the probability that $|\mathcal{K}_{\widehat{\mathcal{C}}} \cap \mathcal{J}| = |\mathcal{J}|$ when $|\mathcal{C}| = c$ given that $|\mathcal{K}_{\widehat{\mathcal{C}}\setminus\{u\}} \cap \mathcal{J}| < |\mathcal{J}|$ and $u \in \mathcal{U}$ is the $\hat{c}^{th}$ user added to $\widehat{\mathcal{C}}$. We condition on the event that $|\mathcal{J}| = \sum_{i=0}^{p-1} |\mathcal{J}_i| = J$. Letting $P_{\mathcal{J}}^i(J_i, c)$ denote the probability that $|\mathcal{J}_i| = J_i$, the probability $P_{\mathcal{J}}(J, c)$ that $|\mathcal{J}| = J$ is equal to the p-fold convolution evaluated at J of the probability distributions $P_{\mathcal{J}}^i(\cdot, c)$. The probability $P_{\mathcal{J}}^i(J_i, c)$ is equal to the summation over all $k \geq J_i$ of the probability that $|\mathcal{K}_{i\mathcal{C}}| = k$ multiplied by the probability that J_i of the k compromised keys are used for jamming, equal to $\binom{k}{J_i}\theta_i^{J_i}(1-\theta_i)^{k-J_i}$. The probability $P_{\mathcal{K}}^i(k, c)$ that $|\mathcal{K}_{i\mathcal{C}}| = k$ is computed recursively by counting the number of new keys recovered from each compromised user. Given that the first $(c - 1)$ compromised users had $(k - n)$ of the q_i keys in $\mathcal{K}_i$, the probability that the m_i keys held by the c^{th} compromised user contain n new keys is $\binom{m_i}{n}\left(1 - \frac{k-n}{q_i}\right)^n \left(\frac{k-n}{q_i}\right)^{m_i-n}$. The remaining probability of interest, denoted $p(\hat{c}|c, J)$, is thus the probability that $|\mathcal{K}_{\widehat{\mathcal{C}}} \cap \mathcal{J}| = J$ when $|\mathcal{C}| = c$ given that $|\mathcal{K}_{\widehat{\mathcal{C}}\setminus\{u\}} \cap \mathcal{J}| < J$ and $|\mathcal{J}| = J$. To compute this probability, we define

$Q_{J,c}(L,\hat{c})$ as the probability that $|\mathcal{K}_{\widehat{\mathcal{C}}} \cap \mathcal{J}| = L$ given $|\mathcal{J}| = J$, $|\mathcal{C}| = c$, and $|\widehat{\mathcal{C}}| = \hat{c}$. If we next condition on the event that $\kappa^{(\hat{c})} = \kappa$, summing over κ with weight $\Phi_{\mathcal{U}}^{(\hat{c})}(\kappa|c) - \Phi_{\mathcal{U}}^{(\hat{c})}(\kappa+1|c)$, this probability can be computed recursively in a similar way to that of $P_{\mathcal{K}}^{i}(k,c)$ above. Given that the first $(\hat{c}-1)$ identified users had $(L-n)$ of the J keys in $\mathcal{J}$, the probability that the κ keys held by the $\hat{c}^{th}$ identified user contain n new keys is $\binom{\kappa}{n}\left(1-\frac{L-n}{J}\right)^{n}\left(\frac{L-n}{J}\right)^{\kappa-n}$. Since the desired probability $p(\hat{c}|c,J)$ is conditional on the event that $|\mathcal{K}_{\widehat{\mathcal{C}}\setminus\{u\}} \cap \mathcal{J}| < J$, and the probability $Q_{J,c}(J,\hat{c})$ is not, conditional probability yields the desired result. □

Proof of Lemma 6: Given $|\mathcal{C}| = c$ and $|\widehat{\mathcal{C}}| = \hat{c}$, the event that F users in $\mathcal{U}\backslash\mathcal{C}$ appear in $\widehat{\mathcal{C}}$ is exactly the intersection of the events $\kappa_{\mathcal{C}}^{(\hat{c}-F)} \geq \kappa_{\mathcal{U}\backslash\mathcal{C}}^{(F+1)}$ and $\kappa_{\mathcal{U}\backslash\mathcal{C}}^{(F)} \geq \kappa_{\mathcal{C}}^{(\hat{c}-F+1)}$. The probability $p(F|\hat{c},c)$ is thus the probability that both of these events occur. The desired probability is evaluated by conditioning on the independent events $\kappa_{\mathcal{U}\backslash\mathcal{C}}^{(F+1)} = \kappa_1$ and $\kappa_{\mathcal{C}}^{(\hat{c}-F+1)} = \kappa_2$. Combining these conditions with the inequalities $\kappa_A^{(n)} \geq \kappa_A^{(n+1)}$ for $A \in \{\mathcal{C}, \mathcal{U}\setminus\mathcal{C}\}$ yields the probability $p(F|\hat{c},c)$ as

$$\begin{aligned} p(F|\hat{c},c) = \sum_{\kappa_1,\kappa_2} & \Pr\left[\kappa_{\mathcal{C}}^{(\hat{c}-F)} \geq \kappa_1 \;\middle|\; \kappa_{\mathcal{C}}^{(\hat{c}-F)} \geq \kappa_2\right] \\ & \times \Pr\left[\kappa_{\mathcal{U}\backslash\mathcal{C}}^{(F)} \geq \kappa_2 \;\middle|\; \kappa_{\mathcal{U}\backslash\mathcal{C}}^{(F)} \geq \kappa_1\right] \\ & \times \Phi_{\mathcal{U}\backslash\mathcal{C}}^{(F+1)}(\kappa_1|c)\Phi_{\mathcal{C}}^{(\hat{c}-F+1)}(\kappa_2|c), \end{aligned} \tag{16}$$

noting that all of the probabilities involved are conditional probabilities given $|\mathcal{C}| = c$. The first probability term in (16) is 1 when $\kappa_2 \geq \kappa_1$ and $\Phi_{\mathcal{C}}^{(\hat{c}-F)}(\kappa_1|c)/\Phi_{\mathcal{C}}^{(\hat{c}-F)}(\kappa_2|c)$ otherwise. Similarly, the second probability term in (16) is 1 when $\kappa_1 \geq \kappa_2$ and $\Phi_{\mathcal{U}\backslash\mathcal{C}}^{(F)}(\kappa_2|c)/\Phi_{\mathcal{U}\backslash\mathcal{C}}^{(F)}(\kappa_1|c)$ otherwise. □

REFERENCES

[1] P. Tague, M. Li, and R. Poovendran, “Probabilistic mitigation of control channel jamming via random key distribution,” in *Proc. 18th Annual IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC’07)*, Athens, Greece, Sep. 2007.

[2] K. Fazel and S. Kaiser, *Multi-Carrier and Spread Spectrum Systems*. Wiley, 2003.

[3] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Prentice Hall, 2001.

[4] J. Schiller, *Mobile Communications*. Addison-Wesley, 2000.

[5] G. L. Stüber, *Principles of Mobile Communications*, 2nd ed. Kluwer, 2001.

[6] R. Anderson, *Security Engineering: A Guide to Building Dependable Distributed Systems*. John Wiley & Sons, Inc., 2001.

[7] A. Chan, X. Liu, G. Noubir, and B. Thapa, “Control channel jamming: Resilience and identification of traitors,” in *Proc. IEEE International Symposium on Information Theory (ISIT’07)*, Nice, France, Jun. 2007.

[8] A. Perrig, J. Stankovic, and D. Wagner, “Security in wireless sensor networks,” *Communications of the ACM*, vol. 47, no. 6, pp. 53–57, Jun. 2004.

[9] W. Xu, W. Trappe, and Y. Zhang, “Channel surfing: Defending wireless sensor networks from interference,” in *Proc. 6th International Conference on Information Processing in Sensor Networks (IPSN’07)*, Cambridge, MA, USA, Apr. 2007, pp. 499–508.

[10] M. Čagalj, S. Čapkun, and J.-P. Hubaux, “Wormhole-based antijamming techniques in sensor networks,” *IEEE Transactions on Mobile Computing*, vol. 6, no. 1, pp. 100–114, Jan. 2007.

[11] L. Eschenauer and V. D. Gligor, “A key-management scheme for distributed sensor networks,” in *Proc. 9th ACM Conference on Computer and Communications Security (CCS’02)*, Washington, DC, USA, Nov. 2002, pp. 41–47.

[12] B. Parno, A. Perrig, and V. Gligor, “Distributed detection of node replication attacks in sensor networks,” in *Proc. 2005 IEEE Symposium on Security and Privacy*, Oakland, CA, USA, May 2005, pp. 49–63.

[13] P. Erdös, P. Frankl, and Z. Füredi, “Families of finite sets in which no set is covered by the union of r others,” *Israel Journal of Mathematics*, vol. 51, no. 1-2, 1985.

[14] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms*. Cambridge: Cambridge University Press, 2003.

[15] R. Diestel, *Graph Theory*, 3rd ed. Springer, 2005.

[16] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, *Handbook of Applied Cryptography*. CRC, 1996.

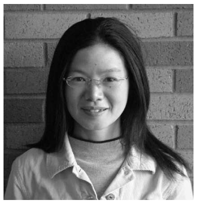
[17] P. Tague and R. Poovendran, “A canonical seed assignment model for key predistribution in wireless sensor networks,” *ACM Transactions on Sensor Networks*, vol. 3, no. 4, Oct. 2007.

[18] ——, “Modeling adaptive node capture attacks in multi-hop wireless networks,” *Ad Hoc Networks*, vol. 5, no. 6, pp. 801–814, Aug. 2007.

[19] W. Feller, *An Introduction to Probability Theory and Its Applications*. John Wiley & Sons, Inc., 1957, vol. 1.

[20] H. A. David and H. N. Nagaraja, *Order Statistics*, 3rd ed. New Jersey: John Wiley & Sons, Inc., 2003.

Patrick Tague is a Ph.D. student in the Electrical Engineering Department at the University of Washington in Seattle. He received his M.S. degree from the same department in 2007 and his B.S. degrees in Mathematics and Computer Engineering from the University of Minnesota, Twin Cities, in 2003. His current research interests include analytical modeling of practical key distribution systems for wireless ad-hoc and sensor networks and attacks and defense mechanisms for distributed wireless networks.



Mingyan Li is an advanced computing technologist in Boeing Phantom Works and an affiliate assistant professor in the department of Electrical Engineering at University of Washington (UW). She received her Doctor of Philosophy degree from Network Security Laboratory at UW in 2006. Her research interests are in the area of network security and user privacy, with applications to sensor networks, RFID systems, software distribution systems, medical security systems, vehicular ad hoc networks (VANET), distributed storage, and secure multicast. Now she is leading Boeing-Siemens collaborative projects on wireless and RFID security. She is a recipient of the departmental Chair’s Award 2006, and the outstanding Society of Women Engineer (SWE) Graduate award 2003.

Radha Poovendran received the Ph.D. degree in Electrical Engineering from the University of Maryland, College Park, in 1999. He is an Associate Professor and founding Director of the Network Security Lab (NSL), Electrical Engineering Department, University of Washington, Seattle. His research interests are in the areas of applied cryptography for multiuser environment, wireless networking, and applications of information theory to security. He is a coeditor of the book *Secure Localization and Time Synchronization in Wireless Ad Hoc and Sensor Networks* (Springer-Verlag, 2007). Dr. Poovendran was a recipient of the NSA Rising Star Award and Faculty Early Career Awards, including the National Science Foundation CAREER Award in 2001, the Army Research Office YIP Award in 2002, the Office of Naval Research YIP Award in 2004, PECASE in 2005 for his research contributions to multiuser security, and a Graduate Mentor Recognition Award from the University of California San Diego in 2006. He co-chaired the first ACM Conference on Wireless Network Security (WiSec) in 2008.