Departmental Papers (ESE)   Department of Electrical & Systems Engineering

4-1-2006

A statistical framework for intrusion detection in ad hoc networks

Dhanant Subhadrabandhu
University of Pennsylvania

Saswati Sarkar
University of Pennsylvania, swati@seas.upenn.edu

Anjun Farooq
Telcordia Technologies, Inc.

A Statistical Framework for Intrusion Detection in Ad Hoc Networks

Dhanant Subhadrabandhu, Saswati Sarkar, Farooq Anjum

Abstract— We focus on detecting intrusions in ad hoc networks using the misuse detection technique. We allow for detection modules that periodically fail to detect attacks and also generate false positives. Combining theories of hypothesis testing and approximation algorithms, we develop a framework to counter different threats while minimizing the resource consumption. We obtain computationally simple optimal rules for aggregating and thereby minimizing the errors in the decisions of the nodes executing the intrusion detection software (IDS) modules. But, we show that the selection of the optimal set of nodes for executing the IDS is an NP-hard problem. We describe a polynomial complexity, distributed selection algorithm, "Maximum Unsatisfied Neighbors in Extended Neighborhood" (MUNEN) that attains the best possible approximation ratio. The aggregation rules and MUNEN can be executed by mobile nodes with limited processing power. The overall framework provides a good balance between complexity and performance for attaining robust intrusion detection in ad hoc networks.

I. INTRODUCTION

Ad hoc networks provide the only means of electronic communication in areas where establishing infrastructure like base stations is either impossible or not cost-effective. Examples include disaster recovery operations, battlefields, communication in remote terrains (e.g., reservations, rural areas), events like superbowl matches, etc. These networks are used by a diverse user population, e.g., civilians in disaster hit areas, spectators in superbowl matches etc., which increases the security risks. One such risk is a user who subverts the functioning of the network by causing undesirable events. Such users are considered as intruders and the events as intrusions. Examples of intrusions are attacks such as TCP SYN flood*, Land Exploit[†], SSPing[‡] etc. [5], [6]. These intrusions leverage system vulnerabilities. There are two ways to prevent such intrusions. One way is to remove the vulnerabilities from the system such as by designing resistant protocols like SCTP [20] to resist TCP SYN flood attacks, patching the operating systems, etc. But, this may not be possible due to various reasons such as poor design [16], limited use of efficient technical solutions (e.g., SCTP is rarely used due to large scale deployment of TCP), different

D. Subhadrabandhu and S. Sarkar are at the Electrical and Systems Engineering Dept of the Univ of Pennsylvania. F. Anjum is at Telcordia Technologies. Their emails are dhanant@seas.upenn.edu, swati@seas.upenn.edu, and fanjum@telcordia.com. The contribution of S. Sarkar was in part supported by the National Science Foundation under grants NCR-0238340 and CNS-0435306.

*The attacker opens a large number of half-open TCP connections.

[†]The attacker sends a TCP SYN packet with the same target and source address.

[‡]The attacker sends a series of fragmented, oversized ICMP packets.

devices having different capabilities, inefficient configuration (e.g., users do not change default security settings or apply patches), etc. The second approach, which is complimentary to the first, is to detect attempts to leverage the vulnerabilities and stop such attempts from succeeding. We focus on the detection aspect of the second approach. We refer to this as intrusion detection.

Intrusion detection has been extensively investigated for wireline networks [7], [10]. But, techniques geared towards wireline networks do not suffice in an ad hoc network due to mobility, the ease of listening to wireless transmissions, lack of fixed infrastructure, etc. [11]. For example, several detection strategies in wireline networks are based on the presence of a small number of static gateways that route and therefore monitor all traffic. But, ad hoc networks typically do not have such choke points, and if such choke-points exist, their locations continuously change due to mobility. Also, intrusions may be detected in wireline networks by detecting anomaly, i.e., by comparing the current system behavior with that in the absence of intrusion. In ad hoc networks, however, normal behavior cannot be accurately characterized, e.g., a node may transmit false updates since the routing protocol is slow to converge and not because it is malicious. Further, unlike in wireline networks, nodes in an ad hoc network have limited energy. Hence, only computationally simple, energy-efficient detection strategies can be used. The detection algorithms must also be distributed as communication with a central computing unit will consume significant energy and bandwidth. Finally, the detection algorithms must seamlessly adapt to topological changes due to mobility. These motivate the design of detection strategies specifically geared towards ad hoc networks.

A strategy specifically suitable for ad hoc networks is that of misuse detection that relies on the use of known patterns of unauthorized behavior. This technique detects intrusion when the transmitted traffic contains abnormal packets which serve as "signatures" of attacks. For example, a UDP packet destined to port 0 can crash some machines [6]. The signature of ping-of-death attack is a very large ping packet, that of RPC locator attack is a packet intended for port 135 that contains a command that the system is not expecting, that of Bubonic attack are various values such as a TTL of 255, a TOS field value of $0xC9$, exactly 20 byte payload in the IP datagram and the fragment ID value with consistent increments of 256 [6]. Due to low false alarm rates, misuse detection is the mainstay of current commercial intrusion detection systems in wireline networks and wireless local area networks. This technique cannot however detect new attacks, i.e., the attacks

whose signatures are unknown. Nevertheless, it is the most suitable technique in ad hoc networks given that it does not require characterization of normal behavior.

But a prerequisite for deploying misuse detection in ad hoc networks is to determine which nodes should execute the sniffing and analysis software modules which we refer to as the intrusion detection system (IDS) modules. A simple strategy is to execute IDS at all nodes [2], [4], [13], [15], [19]. But, executing IDS consumes significant resources like energy, memory and CPU cycles at each node, and nodes have limited resource in a wireless network. Thus, this simple strategy significantly increases the resource consumption in the system. On the other hand, if the IDS are executed in very few nodes, then the resource consumption decreases but several intrusions may escape detection. Huang *et al* [1] propose to organize the network in clusters such that every cluster has a leader and only the leaders monitor the traffic. There is however no guarantee on the resource consumption and the security risk of the above scheme. The challenge is to determine which nodes should execute the IDS so as to minimize the resource consumption subject to limiting the security risk below a tolerable value.

We have recently proposed a framework that attains the above goal in ad hoc networks [24], [25]. The framework however relies on the assumption that the sniffing nodes never generate "false positives", i.e., never conclude that there is an attack when there is none. In practice, however, nodes generate false positives, e.g., when they use bloom filters to detect signatures [8], which in turn significantly complicates the design challenges. In addition, nodes also periodically do not detect signatures in the transmitted traffic even when the transmitted traffic contains the signatures. Now, the security risk of the system must be appropriately formulated so as to consider both missed detections and false positives. Next, the system must attain desired tradeoffs between the above security risk and resource consumption. Last, but not the least, the system needs to decide whether there is an intrusion when different detectors arrive at different decisions, in addition to determining which nodes should execute the IDS. The performance of a decision strategy depend on the selection of the sniffing nodes, and vice-versa, and the security risk and the resource consumption depend on both of these. For example, when large number of nodes execute the IDS, larger number of intrusions are detected as each packet is examined by larger number of nodes, but depending on the decision strategy more false positives may also be generated. Our contribution here is to provide a framework that minimizes the resource consumption subject to limiting the security risk, even when sniffing nodes generate both false positives and missed detections.

We describe our system model and quantify the security risk in Section II. Combining theories of hypothesis testing and approximation algorithms, we develop a framework to counter different errors in decision process, while consuming the minimum possible resource (Section III). We obtain computationally simple optimal and robust rules for aggregating and thereby minimizing the errors in the decisions of the nodes detecting the intrusion. But, we prove that optimally selecting the nodes for sniffing and analyzing packets is NP-hard. We describe a distributed approximation algorithm, "Maximum Unsatisfied Neighbors in Extended Neighborhood" (MUNEN), which attains the best possible approximation ratio while using only simple computations and limited message exchange among nodes. We consider a simple random selection heuristic (RP) which does not involve any message exchange among nodes. Using analysis and simulations, we evaluate the resource consumption and security risks of different decision rules and selection strategies and determine when each may be used (Section IV). We conclude in Section V. We prove the main analytical results in the appendix. Due to lack of space, we prove the remaining analytical results (Theorem 5, Theorem 6, Corollary 1, Corollary 2 and Lemma 1) in [23].

II. SYSTEM MODEL

We first postulate that ad hoc networks in near future will consist of two classes of nodes: (i) nodes that both communicate using the network and perform system tasks like relaying packets, discovering routes, securing communication, etc. (*insider nodes*), and (ii) nodes that only communicate using the network (*outsider nodes*). Our postulate is based on the observation that providing the desired quality of service to users is a pre-requisite for large scale use of this technology. But, if the network is to provide any quality of service (QoS) guarantee it can utilize the users but cannot solely rely on them. This is because users may be available for short durations only. The QoS guarantees can however be provided if some easily deployable low complexity system nodes e.g., static and mobile access points are available. These nodes together with users who are trusted by the network and are in the network most of the time can be relied upon for performing system tasks. Such system nodes and trusted users therefore constitute the insiders. The remaining nodes are the outsiders.

We now provide several example wireless networks that consist of insiders and outsiders. During an event which is widely attended and lasts for short time, e.g., a super-bowl match, service providers may augment the connectivity and coverage provided by the existing cellular and/or Wi-Fi networks by utilizing additional static and mobile access points and the terminals of trusted users [12]. Here, the static and mobile access points as well as the trusted users constitute insider nodes. The remaining users who only communicate using the network are the outsiders. Mesh networks also consist of insiders (mesh points) and outsiders (users). In future, such networks may utilize some trusted users to perform system tasks, particularly during service outage due to failure of existing mesh points, or sudden and temporary increase in service demand in specific areas (temporary hot-spots) - such users would also constitute insiders. Finally, a disaster recovery team can use ad hoc networks to provide services like email, news, audio/video applications etc. in an area where communication infrastructure has been damaged due to a natural disaster or terrorist activity. The insider nodes are access points on buildings and mobile terminals carried by the personnel. The outsider nodes are civilians who communicate using the network.

All the above examples, and more generally the wireless networks with insiders and outsiders, retain the essential char-

acteristics of ad hoc networks. These networks use multi-hop wireless communication, as source-destination paths may involve several insiders who relay messages using wireless links. Nodes in such networks, outsiders and also insiders, may be small mobile terminals and may have limited energy and memory, e.g., access points, laptops, PDAs carried by members of a disaster recovery team and trusted users (insider nodes). Static access points in some existing ad hoc networks in rural areas also have limited energy [3]. Finally, the set of insiders may change with time. For example, the network provider will need to provide incentives in lieu of service to the users who serve as insiders, and hence may utilize such users only as required, e.g., in hot-spots or when existing access points fail. We focus on detecting intrusion in these ad hoc networks. Note that these networks are significantly different from cellular networks where only the last hop is a wireless link, and only the nodes that use the network are mobile, dynamic and have limited energy and memory, while the set of nodes (base-stations) that perform system tasks remain the same, do not change locations and have practically unlimited energy and memory.

We now describe the security risks. An outsider may wish to deliver malicious (*bad*) packets to the destination, which may be an insider or an outsider, resulting in malfunction or failure of the destination. An outsider that sends bad packets is referred to as an intruder. A packet that is not *bad* is referred to as *good*. The network may have multiple intruders. The number and location of the intruders, their destinations and the paths used by them are not known to the network, and vary with time. For simplicity we assume that each attack consists of one packet, e.g., a land exploit attack [6] consists of a bad packet. Note, though, that our approach directly extends to the case where an attack consists of multiple packets [23].

Insider nodes execute the IDS modules that employ misuse-based detection strategy so as to detect bad packets while in transit between the intruder and the destination. We therefore consider the network intrusion detection (NID) technique [14] where the IDS is executed at the network layer of some selected insider nodes. Some insider nodes may not have the capability to execute the IDS. Thus, insider nodes are of two types: (a) *IDS capable* and (b) *IDS incapable*. Also, different IDS capable insider nodes e.g., PDAs, laptops, access points etc. consume different amount of resources to execute the IDS, since they have different residual energy and computational capability. *An IDS capable insider node i has weight w_i that represents its resource consumption when it executes the IDS.* Depending on the system policy, some but not all the IDS capable insider nodes will execute the IDS - these are denoted as *IDS active*.

We represent a wireless network by an undirected graph $G(V, E)$. Here, $V = \{1, \ldots, N\}$ consists of the insiders and E is the set of edges between the insiders. There exists an undirected edge between any two insiders that can receive transmissions from each other. We assume that every insider can receive its own transmission and hence has an edge to itself.

Definition 1: A *neighborhood* N_i of an insider node i is the set of insiders that have edges from i. An insider i *covers* every insider in its neighborhood.

By this definition, an insider is always its own neighbor and covers itself.

An IDS active insider operates in promiscuous mode, i.e., receives any packet that is transmitted by any of its neighbors. An IDS active insider may sporadically fail to detect bad packets and report good packets as bad. A node may not detect bad packets during power saving operations, or if the attack has been designed to evade its IDS module [18], or if the IDS modules on the node are out of date, or if the attacker successfully launches a denial of service (DOS) attack on the node, or if it does not receive the packets due to collisions[§], poor transmission quality in wireless links, etc. Depending on the signature matching techniques used an insider may also report a good packet as bad. Bloom filters [8] are examples of such techniques. Bloom filters consider a packet or packet fragments to be suspicious whenever a hash function of the packet or packet fragments match that of an attack signature. Only suspicious packets are analyzed further. This technique can be implemented in hardware and therefore signatures can be matched at line speed. But, this technique also results in false positives. The false positive rate can sometimes be non-negligible, e.g., sometimes 10% of the good packets have been reported to be suspicious [8]. Now when an insider node is overloaded or does not have enough resource it may report suspicious packets as bad without conducting a thorough analysis. Otherwise, if the node chooses to thoroughly analyze suspicious packets it might have to drop several incoming packets which are more likely to be good. Since at different times different nodes conduct power saving operations or experience poor transmission conditions in different links or are overloaded or suffer DOS, their conclusions about the status of the same packet may be different.

We assume that every IDS active insider considers a bad packet to be good with probability p and a good packet to be bad with probability q. We focus on the efficacy of detection schemes given such a failure model. We do not consider the threats where intruders control the insider nodes. We assume that a packet transmitted by an intruder is good with probability π_G. We consider scenarios where an insider knows π_G (e.g., from history of attacks) and p, q (e.g., from online measurements), and also scenarios where an insider does not know these parameter values.

We assume that either a packet is not encrypted or only the transport layer payload of a packet is encrypted. This is the case with several protocols like PGP, SRTP, HTTPS etc. Then, IDS modules can detect attacks at the transport and lower layers, e.g., ping-of-death, TCP SYN flood, etc. without any knowledge of the encryption schemes. We do not consider link layer and network layer encryption protocols like IPSec since

---

[§]Some collisions happen only due to promiscuous operation. Consider an ad hoc network with 3 insider nodes A, B, C. All nodes are IDS capable. Let both A and C be B's neighbors. But, A and C are not each others neighbors. Let B execute the IDS. If A and C simultaneously relay bad packets to outsider nodes, the packets collide at B. Since however the intruders transmit bad packets only rarely, such collisions of bad packets are rare.
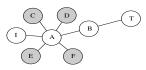
Fig. 1. This figure illustrates the coverage redundancy of an insider. The intruder (I) attacks the destination (T). Insider nodes A and B relay the intruder's packets to the destination. Node A is covered by IDS active insiders (C, D, E, F). When A relays a packet, C, D, E, F receive the packet in promiscuous mode. If any of these detect the packet to be bad, it reports its diagnosis to A. Based on the reports, A determines whether the packet is bad.

these protocols are typically used in enterprise environments which is beyond the scope of this paper. This is a standard assumption in several papers that investigate NID [14].

III. ALGORITHMS FOR ROBUST INTRUSION DETECTION

We first consider the detection goals. Ideally, we would like to maximize the probability of detecting bad packets, and minimize the probability of reporting a good packet as bad ("*false positive*"). But, it may not be possible to attain both goals simultaneously. This happens when increasing the detection rate involves increasing the false positive rate. Thus, our goal is to select the IDS active insiders among the IDS capable insiders so as to minimize the resource consumption subject to limiting the *system risk* below an acceptable value. We quantify the risk as follows. The risk for a bad packet is a constant y_M times the probability P_M that the packet is not detected (*missed detection*). The risk for a good packet is a constant y_F times the probability P_F that the packet is reported as bad. *The expected risk for a packet is therefore* $(1 - \pi_G)y_M P_M + \pi_G y_F P_F$. *We refer to this expected risk as the system risk.* The constants y_M, y_F reflect the relative risks associated with missed detection and false positive. *The resource consumption is the sum of the weights of all IDS active insiders.*

The system risk needs to be maintained below an acceptable value without any knowledge of the intruders, targets and the paths between them. If every IDS active insider could decide without any error ($p = q = 0$), then the expected risk associated with any packet that is analyzed by at least one IDS active insider is 0. Thus, the system risk can clearly be limited if the IDS active insiders are selected so that most packets are analyzed by at least one IDS active insider. Now, since IDS active insiders may decide erroneously, and different IDS active insiders can conclude differently about the status of the same packet, packets must be analyzed by multiple (k or more) IDS active insiders and the results of the analysis intelligently combined to reduce the errors in the decisions. Here, k must be determined in accordance with the acceptable risk.

Due to promiscuous operation, every IDS active insider can receive and hence analyze all packets transmitted by its neighbors. Thus, if the IDS active insiders are selected so that every insider is covered by at least k IDS active insiders, every packet transmitted by an insider would be analyzed by k or more IDS active insiders. Refer to figure 1 for an example detection procedure. Now, all packets, except those that are

directly transmitted from an intruder to its target (which usually constitutes a small fraction of the total number of packets [22]), are relayed by the insiders. Thus, most packets are analyzed by k or more IDS active insiders. *Thus, the system risk can be limited as desired and the resource consumption minimized by optimally (i) aggregating the analysis of different IDS active insiders and (ii) selecting k and subsequently selecting the IDS active insiders such that every insider is covered by at least k IDS active insiders.* We now discuss the issues involved in determining each of the above.

Due to the coverage redundancy several insiders may analyze a packet, and they may decide differently whether the packet is bad. The different decisions must be combined to determine whether the packet is indeed bad. For example, in figure 1, if C and D decide that a packet relayed by A is bad and E and F determine otherwise, A needs to decide whether it should report an attack. The challenge is to aggregate the insiders' decisions so as to minimize the expected risk. Since a packet may be relayed by multiple insiders, the IDS active insiders analyzing a packet may not be geographically close. The aggregation therefore needs to be distributed. *We develop an optimal aggregation scheme based on hypothesis testing framework that minimizes the expected risk and is naturally amenable towards distributed implementation.* An aggregation scheme that is optimal at a given π_G, p, q may lead to a high risk at a different π_G, p, q. Again, *using the hypothesis testing framework, we propose a robust aggregation strategy that depends on p, q but is guaranteed to deliver a certain maximum expected risk irrespective of π_G.* This robustness with respect to π_G is an important advantage as intruders can dynamically vary π_G easily but cannot easily control p, q. Nevertheless, the robustness with respect to all three parameters π_G, p, q can be attained at high k and certain ranges of p, q. *We propose an aggregation strategy that attains a near-zero expected risk when k is high at all π_G, $p \in [0, 0.5), q \in [0, 0.5)$.*

We prove that under the optimal and robust aggregation schemes the expected risk decreases with increase in k. But, the resource consumption also increases with increase in k. The challenge therefore is to select the minimum k that attains a tolerable risk when each insider optimally or robustly aggregates its neighbors' decisions. We quantify the expected risks of different aggregation strategies as functions of k and appropriately select k using these analytical expressions. Once k is determined, we need to select the IDS active insiders so as to minimize the resource consumption or the total weight of the IDS active insiders subject to ensuring that every insider is covered by at least k IDS active insiders. We prove that this problem is NP-hard. Using the theory of approximation algorithms, we design a distributed selection algorithm that attains a guaranteeable approximation bound. The selection algorithms do not depend on the parameters π_G, p, q once k is selected. If robust aggregation rules are used, the selection of k also does not depend on π_G, p, q.

We initially assume that only one insider relays each packet. Under this assumption, in subsection III-A we determine the optimal and robust aggregation rules, and in subsection III-B we determine the optimal value of k and the selection of IDS

active insiders for different aggregation rules. In subsection III-C, we generalize the entire framework to consider arbitrary number of relays between an insider and its target and also relax several other simplifying assumptions made in other subsections.

A. Optimal and Robust aggregation of decisions of different IDS active insiders

We obtain the optimal aggregation scheme under the assumption that each packet is relayed by one insider (different packets can use different relays) and every insider has k IDS active neighbors. Thus, when an insider i relays a packet, its k IDS active neighbors analyze the packet (one of the neighbors may be i itself) and communicate their decisions to i. Then i aggregates its neighbors decisions so as to determine whether the packet is bad.

Definition 2: An aggregation strategy is *optimal* if it minimizes the system risk for a given value of k.

Theorem 1: Let threshold $T_{\text{opt}} = \left\lceil \frac{\ln \frac{y_F \pi_G}{y_M(1-\pi_G)} + k\ln \frac{1-q}{p}}{\ln((1-p)(1-q)/pq)} \right\rceil$. The following is the optimal aggregation strategy for each relay insider.

When $p+q < 1$,[¶] a relay insider decides that a packet is bad if and only if T_{opt} or more of its IDS active neighbors inform that the packet is bad.

When $p+q > 1$, a relay insider decides that a packet is bad if and only if T_{opt} or fewer of its IDS active neighbors inform that the packet is bad.

Let $p + q = 1$. If $\pi_G \geq \frac{y_M}{y_F + y_M}$ a relay insider decides that every packet is good, and otherwise every packet is bad.

Theorem 2: Let $H_{\text{opt}}(k)$ be the minimum system risk.

When $p+q < 1$, $H_{\text{opt}}(k) = y_F \pi_G \sum_{i=\max(T_{\text{opt}},0)}^{k} \binom{k}{i} q^i (1-q)^{k-i} + y_M (1-\pi_G) \sum_{i=0}^{\min(T_{\text{opt}}-1,k)} \binom{k}{i} p^{k-i}(1-p)^i$.

When $p+q > 1$, $H_{\text{opt}}(k) = y_F \pi_G \sum_{i=0}^{\min(T_{\text{opt}},k)} \binom{k}{i} q^i (1-q)^{k-i} + y_M (1-\pi_G) \sum_{i=\max(T_{\text{opt}}+1,0)}^{k} \binom{k}{i} p^{k-i}(1-p)^i$.

When $p+q = 1$, $H_{\text{opt}}(k) = \min (y_M(1-\pi_G), y_F \pi_G)$.

Note that the above aggregation rule is optimal for any relay insider irrespective of whether it is IDS active. Since each insider is also its own neighbor, if a relay insider is IDS active the set of its IDS active neighbors includes itself, and it executes the above aggregation rule considering both its and its other neighbors' analysis of each packet.

We now present the intuition behind the results. When $p + q < 1$, the error probabilities of each IDS active insider's analysis are small for both good and bad packets. So, a large number of insiders are likely to report a packet as bad only when the packet is bad. Thus, a relay insider decides that a packet is bad only when many of its IDS active neighbors report it as bad. Now let $p+q > 1$. Since the error probabilities of each IDS active insider's analysis are high, if a packet is bad (good), many insiders would report it as good (bad). Thus, the previous policy is reversed. This is equivalent to reversing the decision of each IDS active insider and aggregating the reversed decisions

[¶]Note that p and q are probabilities associated with different packets. Thus, $p + q$ can exceed 1.

using the same rule as for $p + q < 1$. Due to the reversal, the effective error probabilities of each insider's analysis are $1 - p$ and $1 - q$ for bad and good packets respectively. Now, the sum of the two new error probabilities $2 - (p+q)$ is less than 1 (as $p + q > 1$) and decreases with further increase in $p+q$. Thus, intuitively the uncertainty in each insider's analysis of a packet decreases with increase in $|p + q - 1|$ and is the maximum when $p+q = 1$. Thus, when $p+q = 1$, the optimum aggregation strategy for a relay insider is to ignore the analysis and decide whether a packet is bad based only on the statistical information π_G about the nature of each packet and weights y_F, y_M. Thus covering every insider with k IDS active insiders is redundant in this case. We later discuss the coverage issues in greater detail (Subsection III-B).

We have so far implicitly assumed that each insider knows π_G, p, q. Note that π_G is the most difficult to ascertain as it is directly controlled by the intruders and they can dynamically vary π_G. The optimum aggregation rule at a given π_G can be substantially suboptimal at a different value of π_G. Thus, the intruders can significantly increase the system risk by selecting a π_G which is different from that assumed by the insiders. The aggregation strategies need to be *robust* to such dynamic variations.

We first quantify a robustness goal. At a given π_G, the expected risk under an aggregation strategy $\mathcal{A}$ is denoted as $R_{\mathcal{A}}(\pi_G)$. The maximum expected risk of $\mathcal{A}$ is the maximum value of $R_{\mathcal{A}}(\pi_G)$ for all possible π_G, i.e., $\max_{0 \leq \pi_G \leq 1} R_{\mathcal{A}}(\pi_G)$. Now, $\max_{0 \leq \pi_G \leq 1} R_{\mathcal{A}}(\pi_G)$ corresponds to the maximum "damage" an intruder can cause by appropriately selecting π_G when the system selects $\mathcal{A}$. The aggregation strategy that minimizes this maximum damage among all aggregation strategies is referred to as a *robust* aggregation strategy. The definition follows.

Definition 3: An aggregation strategy $\mathcal{B}$ is *robust* if for a given value of k, $\max_{0 \leq \pi_G \leq 1} R_{\mathcal{B}}(\pi_G) = \min_{\mathcal{A}} \max_{0 \leq \pi_G \leq 1} R_{\mathcal{A}}(\pi_G)$.

We now present a robust aggregation strategy.

Theorem 3: Let threshold T_{rob} and probability r_{rob} be

$$
T_{\text{rob}} = \begin{cases}
\min i : (y_F \sum_{j=i}^{k} \binom{k}{j} q^j (1-q)^{k-j} \\
\qquad \leq y_M \sum_{j=0}^{i-1} \binom{k}{j} p^{k-j}(1-p)^j) \\
\qquad \text{if } p+q < 1, \\
\max i : (y_F \sum_{j=0}^{i} \binom{k}{j} q^j (1-q)^{k-j} \\
\qquad \leq y_M \sum_{j=i+1}^{k} \binom{k}{j} p^{k-j}(1-p)^j) \\
\qquad \text{if } p+q > 1.
\end{cases}
$$

$r_{\text{rob}} =$

$$
\begin{cases}
\frac{y_M \sum_{j=0}^{T_{\text{rob}}-1} \binom{k}{j} p^{k-j}(1-p)^j - y_F \sum_{j=T_{\text{rob}}}^{k} \binom{k}{j} q^j (1-q)^{k-j}}{\binom{k}{T_{\text{rob}}-1}\left(y_F q^{T_{\text{rob}}-1}(1-q)^{k-T_{\text{rob}}+1} + y_M p^{k-T_{\text{rob}}+1}(1-p)^{T_{\text{rob}}-1}\right)} \\
\qquad \text{if } p+q < 1, \\
\frac{y_M \sum_{j=T_{\text{rob}}+1}^{k} \binom{k}{j} p^{k-j}(1-p)^j - y_F \sum_{j=0}^{T_{\text{rob}}} \binom{k}{j} q^j (1-q)^{k-j}}{\binom{k}{T_{\text{rob}}+1}\left(y_F q^{T_{\text{rob}}+1}(1-q)^{k-T_{\text{rob}}-1} + y_M p^{k-T_{\text{rob}}-1}(1-p)^{T_{\text{rob}}+1}\right)} \\
\qquad \text{if } p+q > 1.
\end{cases}
$$

Clearly, $0 \leq r_{\text{rob}} < 1$.

The following is the robust aggregation strategy for a relay insider.

Let $p + q < 1$. For each packet a relay insider selects a threshold $T = T_{\text{rob}} - 1$ with probability r_{rob} and a threshold $T = T_{\text{rob}}$ with probability $1 - r_{\text{rob}}$. A relay insider decides that the packet is bad if and only if T or more IDS active neighbors inform the relay insider that the packet is bad.

Let $p + q > 1$. For each packet a relay insider selects a threshold $T = T_{\text{rob}} + 1$ with probability r_{rob} and a threshold $T = T_{\text{rob}}$ with probability $1 - r_{\text{rob}}$. A relay insider decides that the packet is bad if and only if T or fewer IDS active neighbors inform the relay insider that the packet is bad.

Let $p + q = 1$. A relay insider decides a packet is bad with probability $\frac{y_M}{y_M + y_F}$, and good with probability $\frac{y_F}{y_M + y_F}$.

We now describe the intuition behind Theorem 3. First, assume that $p + q \neq 1$. Let $r_{\text{rob}} = 0$. Then both the robust and optimum aggregation rules are threshold type. Thus, the intuition behind the robust aggregation rule is the same as that behind the optimum aggregation rule. The rules however select different thresholds. The robust aggregation rule selects threshold T_{rob} so as to equalize the risks associated with missed detection and false positive, i.e., $y_M P_M = y_F P_F$. This leads to an expected risk of $y_M P_M$ irrespective of π_G. Thus an intruder cannot increase the expected risk of the robust aggregation strategy by appropriately selecting π_G - the robust aggregation strategy is "robust" in this sense. Since the thresholds are discrete integers, an insider cannot always equalize the risks by selecting only the thresholds. Now the probability r_{rob} can have any value in the real interval $[0, 1)$. Thus r_{rob} can be selected to equalize the risks by randomizing the decisions. Now, when $p + q = 1$, the uncertainty in the analysis is high. Thus, a robust aggregation rule for a relay insider is to ignore its neighbors' analysis and decide whether a packet is bad based only on the weights y_F, y_M. The robust and the optimum aggregation rules are similar in this case; the difference being that the optimum aggregation rule uses π_G in its decision process. Furthermore, note that for all p, q, the knowledge of π_G allows the optimum aggregation rule to be deterministic, while the lack thereof forces the robust aggregation rule to be randomized.

Theorem 4: Let $H_{\text{rob}}(k)$ be the system risk and $P_M^{\text{rob}}(k)$ be the missed detection probability under the robust aggregation rule. Then $H_{\text{rob}}(k) = y_M P_M^{\text{rob}}(k)$. Also,

when $p + q < 1$, $P_M^{\text{rob}}(k) = \sum_{i=0}^{T_{\text{rob}}-1} \binom{k}{i} p^{k-i}(1-p)^i - r_{\text{rob}} \binom{k}{T_{\text{rob}}-1} p^{k-T_{\text{rob}}+1}(1-p)^{T_{\text{rob}}-1}$,

when $p + q > 1$, $P_M^{\text{rob}}(k) = \sum_{i=T_{\text{rob}}+1}^{k} \binom{k}{i} p^{k-i}(1-p)^i - r_{\text{rob}} \binom{k}{T_{\text{rob}}+1} p^{k-T_{\text{rob}}-1}(1-p)^{T_{\text{rob}}+1}$,

and when $p + q = 1$, $P_M^{\text{rob}}(k) = \frac{y_F}{y_M + y_F}$.

Note that the expected risk of the robust aggregation rule does not depend on π_G, but, exceeds that of the optimal aggregation rule at any given π_G. Intuitively, the difference is the penalty accrued for the insiders' lack of knowledge of π_G.

We still assume that the insiders know p and q. This is justified since the intruders do not directly control p and q and hence cannot alter them easily. Furthermore, the insiders can learn p and q from measurements. Nevertheless, we now propose some heuristic aggregation rules that may be used when the insiders do not know p and q. We first observe the following.

Corollary 1: Let $y_M = y_F$ and $p = q$. When k is odd, $r_{\text{rob}} = 0$. When k is even, $r_{\text{rob}} = 0.5$. In addition, when $p < 0.5$, $T_{\text{rob}} = \lfloor k/2 \rfloor + 1$, and when $p > 0.5$, $T_{\text{rob}} = \lceil k/2 \rceil - 1$.

Corollary 1 suggests that in many cases limited information about p and q is sufficient to obtain the robust aggregation rule. Specifically, when $p = q < 0.5$ the robust aggregation rule determines the nature of a packet based on the decisions of the majority of its IDS active neighbors. Now, usually $p < 0.5$ and $q < 0.5$. This motivates the heuristic "majority aggregation rule" which does not depend on p and q and can therefore be used when an insider does not know these.

We now describe the *majority aggregation rule*. A relay insider decides that a packet is bad (good) if majority of its IDS active neighbors decide that the packet is bad (good). During a tie, the insider decides with probability 0.5 that the packet is bad. Let $H_{\text{maj}}(k)$ be the expected risk under the majority rule.

Lemma 1: Let $\max(p, q) \leq \alpha < 0.5$. Then, $H_{\text{maj}}(k) \leq \max(y_F, y_M) \sum_{i=0}^{\lfloor k/2 \rfloor} \binom{k}{i} \alpha^{k-i}(1-\alpha)^i$.

Note that $H_{\text{maj}}(k)$ depends on π_G, p, q but its upper bound does not depend on exact values of these. The next corollary follows from Theorems 2, 4 and Lemma 1.

Corollary 2: For each π_G, p and q, $H_{\text{opt}}(k)$ and $H_{\text{rob}}(k)$ are monotonically non-increasing functions of k.

If $p + q \neq 1$, $\lim_{k \to \infty} H_{\text{opt}}(k) = \lim_{k \to \infty} H_{\text{rob}}(k) = 0$.

If $\max(p, q) < 0.5$, $H_{\text{maj}}(k)$ is a monotonically decreasing function of k, and $\lim_{k \to \infty} H_{\text{maj}}(k) = 0$.

We first describe the intuition behind this corollary. For both the optimum and robust aggregation rules, an insider can make better decisions if it has more information, i.e., if it hears from more neighbors. Thus, the expected risks of these rules decrease with increase in k and in the limit converge to 0. From the law of large numbers and since $\max(p, q) < 0.5$, as k increases, the probability that $0.5k$ or more (majority) of the IDS active neighbors of a relay insider analyze a packet erroneously decreases and in the limit converges to 0. Thus, the expected risk of the majority rule decreases with increase in k and in the limit converges to 0.

Corollary 2 suggests that if $\max(p, q) < 0.5$, which is usually the case, the majority aggregation rule is near optimum for large k. Thus, the penalty for an insider's lack of knowledge of p and q is that a large k is necessary and hence a large number of insiders need to be IDS active.

B. Selection of IDS active insiders

We now obtain the optimum value of k and the optimum set of IDS active insiders. The former can be obtained in polynomial complexity using the analytical results developed for different aggregation rules in the previous subsection. We prove that the latter is NP-hard. We subsequently describe a polynomial-complexity distributed approximation algorithm, "Maximum Unsatisfied Neighbors in Extended Neighborhood" (MUNEN), which attains the best possible approximation guarantee for the optimal selection problem [21], [25].

We still assume that only one insider relays each packet. When every insider has k IDS active neighbors, each packet has an expected risk of $H(k)$, where $H(k)$ depends on the aggregation rule. We initially assume that each insider knows p and q, but may or may not know π_G. If an insider knows π_G, it uses the optimum aggregation rule; otherwise, it uses the robust aggregation rule. We initially assume that either all insiders know π_G or none knows π_G. In the former case, $H(k) = H_{\text{opt}}(k)$ and in the latter case, $H(k) = H_{\text{rob}}(k)$. We later generalize to the case when only some insiders know π_G. When $p+q = 1$, neither the optimum nor the robust aggregation strategy depends on the analysis of the packets by the IDS active insiders (Theorems 1 and 3). Thus, no insider needs to execute IDS and $k = 0$ suffice. Thus NID is effective only when $p + q \neq 1$, which henceforth we assume.

We now obtain the optimum value of k. Let the tolerable expected risk be γ. We need to select a k such that $H(k) \leq \gamma$. Since the number of IDS active insiders and hence the resource consumption increases with increase in k, the resource consumption is minimized when $k = k_{\min} = \min_k\{k : H(k) \leq \gamma\}$. Since $p+q \neq 1$, $\lim_{k\to\infty} H_{\text{opt}}(k) = \lim_{k\to\infty} H_{\text{rob}}(k) = 0$ (Corollary 2). Thus, for any positive γ, there exists a $k_{\min}$ that attains an expected risk of γ or less, which can be computed using expressions for $H(k)$ given in Theorems 2 and 4.

Now, *the detection goal of attaining the tolerable expected risk subject to minimizing the total weight of the IDS active insiders is satisfied if the IDS active insiders are selected so as to minimize their total weight subject to ensuring that each insider has at least $k_{\min}$ IDS active neighbors.* This ensures that the expected risk remains tolerable and resource consumption is minimized irrespective of the location of the intruders, their targets and paths between them and any statistical distribution of these quantities. We now discuss how to select the IDS active insiders so as to attain this goal.

The optimal set of IDS active insiders can be computed by solving an integer linear program, MRRR$_{\text{IP}}$ (minimize resource consumption subject to attaining the required risk). Let V' be the set of IDS capable insiders. For each insider i, there exists an integer variable x_i. Now, $x_i = 1$ if i is IDS active, and 0 otherwise. Since every insider must be covered by at least $k_{\min}$ IDS active insiders, $\sum_{j \in N_i \cap V'} x_j \geq k_{\min}$. The goal of MRRR$_{\text{IP}}$ is to minimize the weight of the IDS active insiders, i.e., $\sum_{j \in V'} x_j w_j$ subject to these constraints.

(MRRR$_{\text{IP}}$) Minimize: $\sum_{j \in V'} x_j w_j$

subject to

1) $\sum_{j \in N_i \cap V'} x_j \geq k_{\min}, \forall\, i,$
2) $x_i \in \{0, 1\}\ \forall\ i.$

We refer to MRRR$_{\text{IP}}$ as the optimum selection algorithm. The expected risk of the optimum selection algorithm is below γ. This is because some insider nodes will have more than $k_{\min}$ IDS active neighbors, and $H(k)$ decreases with increase in k (Corollary 2). For example, consider a linear array of insider nodes A, B, C, D, E. Every node is a neighbor of its adjacent nodes. If $k_{\min} = 1$, the optimal algorithm will execute IDS in B and D. Thus, C will have 2 IDS active neighbors. This motivates an analysis of the expected risk attained by the
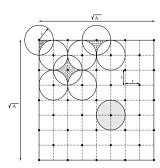

Fig. 2. The figure illustrates an algorithm for selecting the IDS active insiders under the ETICDS(ρ) model. The algorithm would execute IDS in $k_{\min}$ insiders at positions very close to those marked ●. The circles indicate the coverage areas of some of the IDS active insiders.

combination of the optimal selection algorithm and the selected aggregation rule. The analysis will allow us to estimate the difference between γ and the expected risk.

We now describe the first analysis. We assume that N static insiders are uniformly distributed in a square of area A. Each insider is IDS capable and has unit weight. Now we assume that an insider u can receive transmissions from any node v which is within a distance R from u. Thus, the fraction of total area covered by an insider is $\rho = \pi R^2/A$. We refer to this model as the ETICUS(N, ρ) (all insiders have equal transmission ranges, are IDS capable and uniformly distributed in a square) model. Our goal is to compute the expected risk of the optimum selection algorithm $F(N, \rho, k_{\min})$ when a uniformly selected insider v relays the packets and executes the given aggregation rule (optimal or robust). Note that it is difficult to compute $F(N, \rho, k_{\min})$ because of the dependencies between the selection of the IDS active insiders and the topology. So, we approximate this expected risk assuming that every point in the square is an insider. This resembles a network with a large number of insiders. We refer to this assumption as the ETICDS(ρ) model (all insiders have equal transmission ranges, are IDS-capable and densely distributed in a square). Now we execute IDS in insiders at a certain number of locations in the square such that every point in the square is in the coverage area of at least $k_{\min}$ IDS active insiders (figure 2). We now compute the expected risk $F(k_{\min})$ when a uniformly selected point v in the square relays the packets and executes the given aggregation rule (optimum or robust). Now $F(k_{\min})$ represents the expected risk for a selection algorithm and the given aggregation rule, and is therefore expected to be an upper bound for $F(N, \rho, k_{\min})$.

Theorem 5: $F(k_{\min}) = \frac{4-\pi}{2} H(k_{\min}) + \frac{\pi-2}{2} H(2k_{\min})$.
As discussed before, $H(k) = H_{\text{opt}}(k)$ for the optimal strategy and $H(k) = H_{\text{rob}}(k)$ for the robust strategy. These can be computed from Theorems 2 and 4 respectively.

We now examine the computational complexity of the optimal selection algorithm. We no longer limit ourselves to the ETICUS(N, ρ) model, and consider arbitrary distributions for insider nodes and arbitrary edge sets. The complexity of solving MRRR$_{\text{IP}}$ is exponential in N. Furthermore, the

following lemma suggests that the optimal selection problem is not likely to be polynomial complexity computable.

Lemma 2: Optimally selecting the IDS active insiders is NP-hard.

Proof: We first describe the set-multicover problem which is well-known to be NP-hard [9]. There exists a set U with elements $\{u_1, \ldots, u_N\}$ and K subsets of U: $U_1, \ldots, U_K$. The goal is to select the minimum number of subsets such that each element belongs to $k_{\min}$ or more selected subsets. We now show that the set-multicover problem can be solved in polynomial complexity if MRRR$_{\text{IP}}$ can be solved in polynomial complexity. Now, consider a wireless network with IDS capable insiders $v_1, v_2, \ldots, v_K$ and IDS incapable insiders $u_1, \ldots, u_N$. Let $E = \{(v_i, u_j), 1 \leq i \leq K, u_j \in U_i\} \cup \{(v_i, v_j), 1 \leq i \leq K, 1 \leq j \leq K\} \cup \{(u_i, u_i), 1 \leq i \leq N\}$. Here, $V' = \{v_1, \ldots, v_K\}$, $N_{v_i} = V' \cup U_i$, for each $i \in \{1, \ldots, K\}$. Thus v_i covers all insiders in U_i and no insider in $U \setminus U_i$. Let the coverage requirement be $k_{\min}$ for each insider. Clearly, any feasible solution for MRRR$_{\text{IP}}$ selects at least $k_{\min}$ insiders from V' and no insider from $V \setminus V'$. Thus, for any feasible solution, insiders in V' are covered by at least $k_{\min}$ IDS active insiders. Thus, the optimal solution for MRRR$_{\text{IP}}$ is a feasible solution that selects minimum number of insiders from V' subject to ensuring that every insider in $u_1, \ldots, u_N$ is covered with at least $k_{\min}$ selected insiders. Thus, the optimal solution for MRRR$_{\text{IP}}$ provides the optimal solution for the set-multicover problem. ∎

We now consider algorithms for approximating the optimal selection problem. The proof for Lemma 2 demonstrates that the optimal selection problem is an instance of the set-multicover problem. Thus, unless $P = NP$, the best possible approximation ratio$^{\|}$ for this selection problem is $\Omega(\ln N)$ (pp.$112 - 116$, [26]). In our earlier work [21], [25], we have proved that there exists a distributed computationally simple approximation algorithm MUNEN, that not only attains the above approximation ratio ($O(\ln N)$) for the set-multicover problem, but also selects the same set of nodes as the best known centralized approximation algorithm. For completeness, we describe MUNEN here. We first introduce a new notion.

Definition 4: The $k_{\min}$-*priority* of an insider node i is the ratio of the number of its neighbors that are covered by fewer than $k_{\min}$ IDS active insiders to its weight w_i.

Let $h(u, v)$ be the minimum number of hops between insiders u, v. Let $N^2(u)$ be the set of IDS-capable $2-$hop neighbors of u not including u, i.e., $N^2(u) = \{v : v \in V' \setminus \{u\}, h(u, v) \leq 2\}$.

In each iteration an IDS-capable insider u selects itself if
- u has not selected itself already and
- u's $k_{\min}$-priority is positive and
- for any other insider v that has not been selected yet and is in $N^2(u)$, either (i) u's $k_{\min}$-priority is greater than that of v or (ii) u's $k_{\min}$-priority equals that of v and $u < v$.

At the end of each iteration, insiders recompute their $k_{\min}$-priorities.

$^{\|}$The approximation ratio of a node selection algorithm is the ratio between the number of nodes it selects and the number of nodes the optimal algorithm selects.

Clearly, MUNEN selects additional IDS active insiders in each iteration until each IDS capable insider that has not been selected has 0 $k_{\min}$-priority. It can select at most $|V'|$ insiders. Thus, MUNEN terminates in $|V'|$ iterations, and the complexity of the computations at each insider is $O(|V'|^2)$.

When an insider node in G is covered by at most $k_{\min} - 1$ IDS capable insiders, it may not be possible to select IDS active insiders so as to cover every insider by $k_{\min}$ IDS active insiders. In this case, MUNEN selects IDS active insiders such that either an insider has at least $k_{\min}$ IDS active neighbors or all its IDS capable neighbors are IDS active. Thus MUNEN provides the maximum possible coverage in this case.

MUNEN is oblivious to the position of the outsiders, and is therefore not affected by their movements. But, it needs to recompute the IDS active set each time an insider node's neighborhood changes due to its or its neighbors' movements. Even though insiders require only local information and limited message exchanges during such computations, frequent execution of these computations may consume significant resource. Thus, MUNEN may not be suitable when the insider nodes move rapidly.

We therefore consider a naive algorithm, Random Placement (RP), in which nodes do not exchange any control message. Here, every IDS capable insider node executes the IDS with a probability s which can be selected so as to regulate the resource consumed and the detection rate. For example, if s is high, then a large number of insiders are IDS active. Thus, the detection consumes a lot of resource but the expected risk is low. We now compute the expected risk obtained by the combination of RP and a selected aggregation rule. Here we restrict ourselves to ETICUS(N, ρ) model. We compute the expected risk $G(N, s, \rho)$ when a uniformly selected insider v relays the packets and executes the given aggregation rule (optimal or robust).

We first compute the probability l_k that an insider has k IDS active neighbors.

$$f(x) = (1 - \frac{\cos^{-1}(x)}{\pi} + \frac{x}{\pi}\sqrt{1 - x^2})$$

$$g(x, y)$$
$$= 1 - \frac{\cos^{-1}(x) - x\sqrt{1 - x^2} + \cos^{-1}(y) - y\sqrt{1 - y^2}}{\pi}$$
$$- \frac{0.5(\sqrt{1 - y^2} - x)(\sqrt{1 - x^2} - y)}{\pi}$$
$$+ \frac{|\pi/4 - (\cos^{-1}(x) + \cos^{-1}(y))/2|}{\pi}$$
$$- \frac{0.5\sqrt{1 - x^2 - y^2 + 2x^2y^2 - 2xy\sqrt{1 - x^2}\sqrt{1 - y^2}}}{\pi}$$

$$l_k = (1 + \frac{4\rho}{\pi} - 4\sqrt{\frac{\rho}{\pi}})(s\rho)^k(1 - s\rho)^{N-k} \times$$
$$\left[\binom{N-1}{k}\frac{1-s}{1-s\rho} + \binom{N-1}{k-1}\frac{1}{\rho}\right]$$
$$+ (4\sqrt{\frac{\rho}{\pi}} - 8\frac{\rho}{\pi})\left[\binom{N-1}{k}(1-s) \times\right.$$
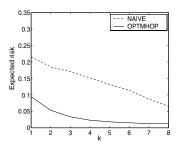
Fig. 3. We now illustrate the importance of intelligent aggregation rules when multiple IDS active insiders relay a packet. We plot the expected risk as a function of k. We present averages over 300 topologies. Each topology has one mobile intruder, one target and $p = 0.4, q = 0.2$. Here we consider ETICUS$(100, 0.087)$ model. The path between an insider and its target, which is selected by AODV, may consist of multiple links. Here, $y_F = y_M = 1$ and $\pi_G = 0.8$. When no insider is IDS active, and every packet is considered as good, the expected risk is 0.2. When IDS active insiders are selected using MUNEN and their analysis aggregated using the intuitively appealing algorithm (NAIVE) the expected risk exceeds 0.2!. The expected risk associated with the optimum multihop aggregation scheme (OPTMHOP) is smaller.

$$\int_0^1 (s\rho f(x))^k (1 - s\rho f(x))^{N-k-1} dx$$
$$+ \binom{N-1}{k-1} s \int_0^1 (s\rho f(x))^{k-1} (1 - s\rho f(x))^{N-k} dx \Bigg]$$
$$+ \frac{4\rho}{\pi} \left[\binom{N-1}{k}(1-s) \int_0^1 \int_0^1 (s\rho g(x,y))^k \times \right.$$
$$(1 - s\rho g(x,y))^{N-k-1} dxdy + \binom{N-1}{k-1} s \times$$
$$\left. \int_0^1 \int_0^1 (s\rho g(x,y))^{k-1} (1 - s\rho g(x,y))^{N-k} dxdy \right]$$

Theorem 6: $G(N, s, \rho) = \sum_{i=0}^{N} l_i H(i)$.

In Section IV, we compare the performances of MUNEN and RP, and determine when each may be deployed.

We have so far assumed that every insider knows p, q. If not, when $\max(p, q) < 0.5$ and an upper bound α of $\max(p, q)$ is known, similar results can be obtained by using the majority aggregation rule. Now, α can often be determined from system knowledge. The upper bound of $H_{\mathrm{maj}}(k)$ provided in Lemma 1 must be used to compute $k_{\min}$. Thus a large $k_{\min}$ would normally be required to reduce this upper bound below γ and hence the resource consumption would be high. This is the penalty for insiders' lack of knowledge of p, q.

C. Generalizations and Discussions

We first generalize the framework to consider arbitrary number of relay insiders between an intruder and its target. Now, the IDS active neighbors of all relay insiders analyze a packet. The main challenge is to aggregate in a distributed manner all this analysis.

We first demonstrate that an intuitively appealing aggregation scheme can lead to a very high expected risk. Consider a multihop aggregation scheme (NAIVE) in which a packet is considered malicious if at least one relay insider considers it such. Each relay insider decides whether the packet is bad using

a single-hop aggregation rule (presented in subsection III-A) based on the analysis of its IDS active neighbors. Clearly, this scheme has high false positives, which leads to significantly high expected risk. Sometimes, the resulting expected risk is above that incurred when no insider is IDS active (fig. 3).

Having shown the need for intelligent aggregation rules, we next describe the optimum and robust aggregation rules for multihop sessions. Irrespective of the number of relay insiders an IDS active insider covers, it analyzes a packet only once. Let k_{mhop}^i be the total number of distinct IDS active neighbors of the relay insiders of the ith session. Now, Theorems 1 and 3 respectively provide the optimal and the robust aggregation rules for each session i with k substituted by k_{mhop}^i and the "IDS active neighbors" referring to the "distinct IDS active neighbors of all relay insiders". The aggregation procedure is therefore similar to that of a fictitious system where every packet of session i is relayed by a single insider with k_{mhop}^i active neighbors.

Since a packet is examined by a larger number of IDS active insiders as the number of hops in its path increases, the expected risks for the above aggregation rules decrease with increase in the path length of the sessions for each selection rule (MUNEN, RP). Thus, selecting $k_{\min}$ as before (i.e., assuming that every packet is relayed by only one insider) still ensures that the expected risk does not exceed γ. Such a selection also ensures that $k_{\min}$ and hence the selection of the IDS active insiders do not change with change in paths between intruders and their targets. Thus, the selection algorithms remain the same irrespective of the path lengths of the sessions.

We now describe a distributed implementation of the above aggregation rules. For simplicity assume that $p + q < 1$. The case of $p+q > 1$ can be treated similarly. Both the optimum and the robust aggregation rules decide whether a packet is good or bad depending on whether the number of IDS active insiders that conclude that the packet is bad exceeds a certain threshold. The last relay insider for session i determines this threshold using k_{mhop}^i and Theorems 1 and 3. Note that k_{mhop}^i can be obtained during route discovery. Now, every insider relaying a packet determines how many of its IDS active neighbors decide that the packet is bad ("bad votes"), and communicates this number to the next relay after relaying the packet. Subsequent relay insiders add the number of bad votes they count to the number sent by the previous relay and transmit the sum further downstream. Thus the last relay insider knows the total number of bad votes. The last relay can thereby detect the occurrence of an attack if this number exceeds the threshold.

Several optimizations are possible. First, the number of bad votes can be communicated in the header of the next packet. This eliminates the need for separate control packet transmission, except once every time a new path is used. Note that several data packets are transmitted between consecutive path changes. Also, every relay insider for session i can determine the threshold for the aggregation rules using k_{mhop}^i and Theorems 1 and 3. Thus, a relay insider can instruct its next hop relay to signal an attack if the number of bad votes obtained so far exceeds this threshold. The integrity of these communications among the insiders may be protected using

message authentication codes - this prevents intruders from posing as insiders.

We now consider some other generalizations of the framework. First, some insiders may know π_G while others may not. Then, the former class of insiders uses the optimal aggregation strategy and the latter uses robust aggregation strategy. Now, at any π_G and k, the expected risk of the optimal aggregation strategy is less than that of the robust aggregation strategy. Thus, the insiders who do not know π_G must be covered by more IDS active insiders for providing the same expected risk for the relayed packets. Specifically, an insider i must be covered by at least k_i IDS active insiders, where $k_i = k_{\min}^{\mathrm{opt}} = \min_k\{k : H_{\mathrm{opt}}(k) \leq \gamma\}$ ($k_i = k_{\min}^{\mathrm{rob}} = \min_k\{k : H_{\mathrm{rob}}(k) \leq \gamma\}$) if i knows (does not know) π_G. Now, clearly the optimal selection problem remains NP-hard. A generalization of MUNEN which considers different coverage requirements for different insiders attains the best possible approximation ratio of $O(\ln N)$.

We have assumed that all the IDS active insiders have equal p, q. All the results hold when all IDS active neighbors of each insider have the same p, q. Now, different insiders have different coverage requirements which depend on the p, q of their IDS active neighbors. This can be accommodated as described in the previous paragraph. The aggregation rules can be generalized to accommodate the more general case of different p, q of different IDS active neighbors of an insider, and the IDS active insiders can now be selected to satisfy the different coverage requirements.

IV. PERFORMANCE EVALUATION

Using ns2-simulations, we compare the performance of different aggregation and IDS active insider selection algorithms. The simulations allow us to investigate the effect of the factors we did not consider in the analysis such as arbitrary number of hops between the intruder and the target, mobile insiders etc. We also evaluate the advantages and disadvantages of different aggregation rules and the benefits of intelligently selecting the IDS active insiders. We accordingly decide the appropriate aggregation rule and the selection algorithm for different desired tradeoffs between tolerable risks and resource consumption.

We consider networks with different p, q ($p = q = 0.1$, $p = 0.3, q = 0.1$) and different types of node mobility. For each combination, we measure averages over 300 different topologies. Each topology consists of a single intruder, a single target. Here we consider ETICUS$(400, 0.05)$ model. Also, $y_M = y_F = 1$. For each topology, we measure the expected risk as the sum of the bad packets that are not detected and the good packets that are reported as bad divided by the total number of packets.

We measure the detection cost as the total number of IDS active insiders. We could not simulate the optimal node selection algorithm MRRR$_{\mathrm{IP}}$ in these large networks due to the computational complexity involved in solving integer linear programs with 400 variables. Nevertheless, sample computations in smaller networks suggest that MUNEN closely approximates MRRR$_{\mathrm{IP}}$, and the performance difference is generally

much less than the upper bound of $O(\ln N)$. We do not include these comparisons due to space constraint.

We initially consider networks where every packet is relayed by a single insider node. In each topology, we select the IDS active insiders using MUNEN and RP. Then, we select an insider uniformly among all the insiders, and measure the risk when it relays the packets and uses different aggregation rules. The intruder and its target are selected within the transmission range of this insider. Here, we assume that all nodes are static.

We investigate the performance of the various aggregation strategies for two selection algorithms MUNEN and RP. In figure 4 we plot the expected risk as a function of π_G. For MUNEN, we consider $k_{\min} = 5$. For RP we select the IDS activation probability s such that both RP and MUNEN have equal number of IDS active insiders. This ensures that both selection algorithms consume equal resource. We plot both the expected risk measured using simulations and that computed using the analytical expressions in Theorems 5, 6, 2, 4. We first compare the performance of different aggregation rules for each selection algorithm, and thereafter comment on the efficacy of the analysis. The figure demonstrates that as expected the optimal aggregation strategy (labeled as OPTAGG) has the minimum expected risk at each π_G. The robust strategy (ROBAGG) on the other hand has somewhat higher risk at all π_G, but this risk does not depend on π_G which again follows from Theorem 4. Also, when $p = q$ (figures 4(a) and 4(c)), the robust and majority aggregation rules have equal risks. This is because the two rules are the same whenever $p = q$ (Corollary 1). But, when $p > q$ (figures 4(b) and 4(d)) and π_G is low the majority aggregation rule has significantly higher expected risk than the other aggregation rules. The expected risk of the majority aggregation rule decreases with increase in π_G. This is because when $p > q$, $P_F < P_M$ for the majority aggregation rule; thus, its expected risk $\pi_G(P_F - P_M) + P_M$ is a monotonically decreasing function of π_G.

Figure 4 demonstrates a close match between the analytical results and simulation measurements for RP– the respective curves are indistinguishable. This is because Theorem 6 provides exact expressions for RP. But, the analytical results for the optimum selection algorithm (Theorem 5) upper bound the simulation measurements for MUNEN. This is again expected as the expected risk in Theorem 5 upper bounds the expected risk of the optimal selection algorithm whose performance is similar to that of MUNEN. Note that the analytical results in Theorem 5 have the same trend as the simulation measurements for MUNEN. The analytical expressions in both Theorem 5 and Theorem 6 are computationally simple, whereas the simulations are computationally intensive. This renders the exact analysis for RP and the approximate analysis for the optimum selection algorithm very useful.

Let the optimal aggregation strategy at a given π_G be denoted as $\mathcal{O}(\pi_G)$. Here, we consider the same $k_{\min}$ and s as before. In figure 5(a), we plot the maximum expected risk of $\mathcal{O}(\pi_G)$, i.e., $\max_{0 \leq x \leq 1} R_{\mathcal{O}(\pi_G)}(x)$ obtained using the expression in Theorem 2, 5, as a function of π_G when $k = 5$. Recall that this maximum expected risk is a measure of the maximum damage an intruder inflicts if each insider tries to minimize the
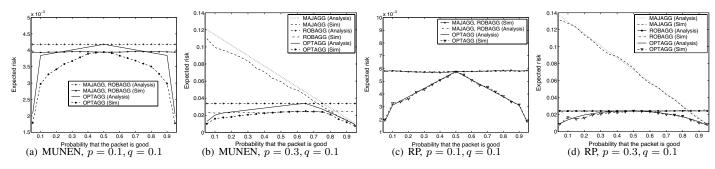
Fig. 4. We plot the expected risk as a function of π_G. Each packet is relayed by a single uniformly selected insider. We obtain the expected risk from both analysis and simulations.
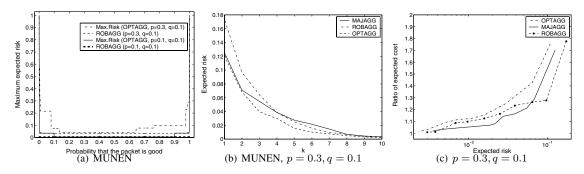

Fig. 5. In figure a), we plot the maximum expected risk as a function of π_G at $k = 5$. In figure b) we consider the expected risk as a function of k for different aggregation strategies. We consider $\pi_G = 0.8$. In figure c), we plot the ratio of expected cost of RP and MUNEN for different aggregation strategies as a function of the expected risk. In all cases, we assume that the packet is relayed by a single uniformly selected insider. We obtain the data in figures a) from the analysis and in figures b) and c) from simulations.

system risk assuming a specific value of π_G and the intruder selects a different π_G. We also plot the maximum expected risk of the robust strategy as a benchmark which as expected does not change with π_G (Theorem 4). The figure demonstrates that the maximum expected risk of the optimal strategy is significantly higher than that of the robust strategy. Moreover, the maximum expected risk is the maximum at extremes of π_G. This is because when the insiders assume one extreme for π_G, the intruder can significantly increase the risk by selecting the other extreme for π_G. *We conclude that the robust aggregation strategy must be used when the intruders select π_G after observing the insiders' aggregation strategy, whereas the optimum aggregation strategy should be used if the intruders do not vary π_G.*

We now examine when the majority aggregation rule can be used. Note that we have so far observed that for $k = 5$ the majority aggregation rule performs significantly worse than the optimum and robust aggregation rules when $p \neq q$. But, again neither the optimum nor the robust aggregation rule can be used when p, q are unknown. Furthermore, Corollary 2 suggests that the majority aggregation rule can be used for large k and any p, q when $\max(p, q) < 0.5$. In figure 5(b), we plot the expected risk as a function of k for different aggregation rules at $p = 0.3, q = 0.1$ and $\pi_G = 0.8$. We select the IDS active insiders using MUNEN. As k increases, the expected risk associated with the various strategies decrease and approach 0. Further, for

large k the expected risks are similar for different aggregation rules. This indicates that *the majority rule can be used with a large k when p, q are unknown.*

We now compare the resource consumed by different IDS active insider selection algorithms for different aggregation strategies. In figure 5(c), we plot the ratio of the expected detection costs of RP and MUNEN measured from simulations as a function of the tolerable expected risk. For MUNEN at each tolerable risk value, we determine $k_{\min}$ and subsequently the expected detection cost at $k = k_{\min}$. For RP, we determine the minimum s required to attain the tolerable expected risk. We now compare RP and MUNEN for different aggregation rules. We first explain the trends and subsequently draw conclusions. When the security requirements are very stringent, i.e., the tolerable expected risk is very low (e.g., less than 0.01), MUNEN and RP have similar detection costs. Now, $k_{\min}$ and s are large in this case. Thus, most of the neighbors of the insiders need to execute the IDS. Hence, both RP and MUNEN would need to execute the IDS in a large fraction of insider nodes leading to similar detection costs. But, when the security requirements are slightly less stringent, i.e., the tolerable risk values are slightly higher, RP consumes significantly higher resource than MUNEN for all aggregation rules. For example, for OPTAGG, for tolerable expected risk values greater than 0.01 and 0.04, RP consumes 12% and 32% more resource than MUNEN respectively. The difference
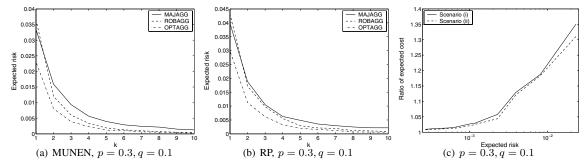
Fig. 6. Multihop sessions with static and mobile insider nodes. We consider topologies where the path between the intruder and the destination consists of multiple hops. In figures a) and b), we plot the expected risk as a function of k for different aggregation strategies for MUNEN and RP respectively. We consider $\pi_G = 0.8$. The insiders are static. In figure c), we plot the ratio of expected cost of RP and MUNEN as a function of the expected risk for optimal aggregation strategy. In all cases, we obtain the data from simulations.

between the resource consumed becomes much higher for larger values of tolerable risks. In this region, intermediate values of $k_{\min}$ and s are required, and MUNEN's intelligent selection of IDS active insiders attains the same coverage as RP while using fewer IDS active insiders. Summarizing, *MUNEN significantly outperforms RP for most values of tolerable risks.*

The ratio of the detection costs of RP and MUNEN is somewhat higher for the optimum aggregation rule than other aggregation rules as the former requires somewhat lower $k_{\min}$ and s at any given value of tolerable expected risk. Thus, as discussed before, MUNEN's intelligent selection of IDS active insiders is more effective for the optimum aggregation rule.

We next consider topologies with an arbitrary number of relays between an intruder and its target. The intruder and its target are selected uniformly. Thus the path between them, which is selected by AODV, consists of arbitrary number of hops. We consider two different scenarios: (i) static insiders and a mobile intruder and (ii) mobile insiders and a mobile intruder. Each mobile node moves as per the random way point model with a maximum speed of 20 m/s and pause time 10 sec. In figures 6(a) and 6(b), we plot the expected risk as a function of k for different aggregation strategies for scenario (i). In figure 6(c), we plot the ratio of the expected detection costs of RP and MUNEN as a function of the tolerable expected risk. We present results for both scenarios (i) and (ii) in this figure considering only the optimal aggregation rule. In both these figures, the trends remain the same as for a single hop network. Thus our conclusions remain the same.

V. CONCLUSION

We consider ad hoc networks with imperfections in the nodes performing intrusion detection tasks. Combining tools from the theories of hypothesis testing and approximation algorithms, we develop a framework to counter different threats while minimizing the resource consumption. We obtain computationally simple optimal and robust rules for aggregating and thereby minimizing the errors in the decisions of the nodes detecting the intrusion. But, we show that optimally selecting the nodes for sniffing and analyzing packets has the same complexity as the well-known NP-hard set-multicover problem

[9]. We next describe a polynomial-complexity distributed approximation algorithm, MUNEN, for selecting the nodes for sniffing and analyzing packets which attains the best possible approximation ratio and a simple heuristic selection strategy (RP). Finally, we evaluate the security risk and the resource consumed by the decision rules and the selection strategies using both analysis and simulations. The overall framework provides a good balance between complexity and performance for detecting intrusion in ad hoc networks.

APPENDIX

Proof of Theorem 1: Let H_0 (H_1) denotes the hypothesis that the packet is good (bad). Thus, when a relay insider decides that a packet is good (bad), it accepts hypothesis H_0 (H_1). Also, missed detection (fault alarm) corresponds to accepting H_0 (H_1) when H_1 (H_0) holds. Let the cost for accepting H_0 (H_1) when H_0 (H_1) does not hold be y_M (y_F). Thus the decision rule that minimizes the expected risk of a relay insider's decision also minimizes the expected cost of accepting a hypothesis. Now let $P_0(y)$ $(P_1(y))$ be the probability that y IDS active neighbors of a relay insider inform that the packet is bad given that H_0 (H_1) holds. Then, $P_0(y) = \binom{k}{y}q^y(1-q)^{k-y}$ and $P_1(y) = \binom{k}{y}(1-p)^y p^{k-y}$.

Now from decision theory [17] (pp.5-9), the decision rule that minimizes the expected cost of accepting a hypothesis (H_0 or H_1) is the following. A relay insider selects the hypothesis H_1 if $\frac{P_1(y)}{P_0(y)} \geq \frac{\pi_G y_F}{(1-\pi_G)y_M}$ and selects the hypothesis H_0 otherwise. By replacing the expressions for $P_1(y)$ and $P_0(y)$, the optimal decision rules given in Theorem 1 for the cases when $p + q < 1$ and $p + q > 1$ follow. Refer to [23] for the case when $p + q = 1$.

Proof of Theorem 2: Let $P_F^{\mathrm{opt}}(k)$ and $P_M^{\mathrm{opt}}(k)$ denote the probability of false positive and missed detection under the optimal decision rule respectively. From Theorem 1, when $p + q < 1$, $P_F^{\mathrm{opt}}(k) = \sum_{i=\max(T_{\mathrm{opt}},0)}^{k} \binom{k}{i}q^i(1-q)^{k-i}$ and $P_M^{\mathrm{opt}}(k) = \sum_{i=0}^{\min(T_{\mathrm{opt}}-1,k)} \binom{k}{i}p^{k-i}(1-p)^i$. When $p + q > 1$, $P_F^{\mathrm{opt}}(k) = \sum_{i=0}^{\min(T_{\mathrm{opt}},k)} \binom{k}{i}q^i(1-q)^{k-i}$ and $P_M^{\mathrm{opt}}(k) = \sum_{i=\max(T_{\mathrm{opt}}+1,0)}^{k} \binom{k}{i}p^{k-i}(1-p)^i$. Since $H_{\mathrm{opt}}(k) = (1 -$

$\pi_G)y_M P_M^{\text{opt}}(k) + \pi_G y_F P_F^{\text{opt}}(k)$, the results follow. Refer to [23] for the proof when $p + q = 1$.

Proof of Theorem 3: Let $\delta_{\pi_G}(k)$ be the optimal decision rule of a relay insider obtained at a given value of π_G and k. Now consider the case when $p + q < 1$, for given values of p and q. Let 1) $R_M(\delta_{\pi_G}(k)) = y_M \sum_{j=0}^{\min(T_{\text{opt}}-1,k)} \binom{k}{j} p^{k-j}(1-p)^j$ and 2) $R_F(\delta_{\pi_G}(k)) = y_F \sum_{j=\max(T_{\text{opt}},0)}^{k} \binom{k}{j} q^j(1-q)^{k-j}$. Let the expected risk incurred by the optimal decision rule $\delta_{\pi_G}(k)$ be $H(\delta_{\pi_G}(k))$. Hence, $H(\delta_{\pi_G}(k)) = \pi_G R_F(\delta_{\pi_G}(k)) + (1-\pi_G)R_M(\delta_{\pi_G}(k))$. $H(\delta_{\pi_G}(k))$ can be rewritten as $H(\delta_{\pi_G}(k)) = [R_F(\delta_{\pi_G}(k)) - R_M(\delta_{\pi_G}(k))]\pi_G + R_M(\delta_{\pi_G}(k))$. Now we want to show that $H(\delta_{\pi_G}(k))$ is a piecewise linear function of π_G. From Theorem 1, T_{opt} is a nondecreasing step function of π_G. In addition, $R_M(\delta_{\pi_G}(k))$ and $R_F(\delta_{\pi_G}(k))$ depend on π_G only through T_{opt}. Hence, $R_F(\delta_{\pi_G}(k))(R_M(\delta_{\pi_G}(k)))$ is a non-increasing (non-decreasing) function of T_{opt}. Thus, $R_F(\delta_{\pi_G}(k))$ and $R_M(\delta_{\pi_G}(k))$ are also step functions of π_G. Hence, the slope of $H(\delta_{\pi_G}(k))$ is a non-increasing step function of π_G. Thus, $H(\delta_{\pi_G}(k))$ is a piecewise linear function. It is also straightforward to show that $H(\delta_{\pi_G}(k))$ is a continuous concave function [17] (pp.14). Now let $\pi_{G'}$ be the minimum value of π_G such that $H(\delta_{\pi_G}(k))$ is maximized. Let $T_1(T_2) = \lim_{\pi_G \uparrow \pi_{G'}} T_{\text{opt}}(\pi_G) (\lim_{\pi_G \downarrow \pi_{G'}} T_{\text{opt}}(\pi_G))$. Note that $T_1(T_2)$ is the threshold associated with δ_{π_G} for some $\pi_{G^1}(\pi_{G^2})$ where $\pi_{G^1} < \pi_{G'}(\pi_{G^2} > \pi_{G'})$. Since $H(\delta_{\pi_G}(k))$ is concave and $H(\delta_{\pi_{G'}})$ is the global maximum, it is straightforward to see that $R_F(\delta_{\pi_{G^1}}(k)) - R_M(\delta_{\pi_{G^1}}(k)) > 0$ while $R_F(\delta_{\pi_{G^2}}(k)) - R_M(\delta_{\pi_{G^2}}(k)) \le 0$. Also, $T_2 = T_1 + 1$. Thus, $T_2 = \min i : y_F \sum_{j=i}^{k} \binom{k}{j} q^j(1-q)^{k-j} \le y_M \sum_{j=0}^{i-1} \binom{k}{j} p^{k-j}(1-p)^j$.

Now from decision theory [17] (pp.13-18), the decision rule that minimizes the maximum expected risk among all decision rules is the following. *A relay insider uses the decision rule* $\delta_{\pi_{G^1}}(k)$ *with probability* r_{rob} *and uses the decision rule* $\delta_{\pi_{G^2}}(k)$ *with probability* $1 - r_{rob}$ where
$r_{\text{rob}} = \frac{R_F(\delta_{\pi_{G^2}}(k)) - R_M(\delta_{\pi_{G^2}}(k))}{R_F(\delta_{\pi_{G^2}}(k)) - R_M(\delta_{\pi_{G^2}}(k)) + R_M(\delta_{\pi_{G^1}}(k)) - R_F(\delta_{\pi_{G^1}}(k))}$.
Recall that a relay insider which uses $\delta_{\pi_{G^1}}(k)$ $(\delta_{\pi_{G^2}}(k))$ decides that the packet is bad if and only if $T_1(T_2)$ or more of its IDS active neighbors inform that the packet is bad. Now, by replacing the expressions for $R_F(\delta_{\pi_{G^2}}(k)), R_M(\delta_{\pi_{G^2}}(k)), R_F(\delta_{\pi_{G^1}}(k)), R_M(\delta_{\pi_{G^1}}(k)), T_1$ and T_2, the results follow. Refer to [23] for the cases when $p + q > 1$ and $p + q = 1$.

Proof of Theorem 4: Let $P_F^{\text{rob}}(k)$ and $P_M^{\text{rob}}(k)$ denote the probability of false positive and missed detection under the robust decision rule respectively. From Theorem 3, when $p + q < 1$, $P_M^{\text{rob}}(k) = r_{\text{rob}} \sum_{i=0}^{T_{\text{rob}}-2} \binom{k}{i} p^{k-i}(1-p)^i + (1-r_{\text{rob}}) \sum_{i=0}^{T_{\text{rob}}-1} \binom{k}{i} p^{k-i}(1-p)^i$ and $P_F^{\text{rob}}(k) = r_{\text{rob}} \sum_{i=T_{\text{rob}}-1}^{k} \binom{k}{i} q^i(1-q)^{k-i} + (1-r_{\text{rob}}) \sum_{i=T_{\text{rob}}}^{k} \binom{k}{i} q^i(1-q)^{k-i}$. Recall that $H_{\text{rob}}(k) = (1-\pi_G)y_M P_M^{\text{rob}}(k) + \pi_G y_F P_F^{\text{rob}}(k)$. By substituting the expression for r_{rob} in $P_F^{\text{rob}}(k)$ and $P_M^{\text{rob}}(k)$, we have $y_M P_M^{\text{rob}}(k) = y_F P_F^{\text{rob}}(k)$. By replacing $y_F P_F^{\text{rob}}(k)$ by $y_M P_M^{\text{rob}}(k)$ in $H_{\text{rob}}(k)$, the result follows. Refer to [23] for the cases when $p + q > 1$ and $p + q = 1$.

REFERENCES

[1] Yi an Huang and Wenke Lee. A cooperative intrusion detection system for ad hoc networks. *Proceedings of the ACM Workshop on Security of Ad Hoc and Sensor Networks (SASN'03)*, Oct 2003.

[2] F. Anjum and R. Talpade. Packet-drop detection algorithm for ad hoc networks. In *Proc. of 60th IEEE Vehicular Technology Conference*, Sept. 2004.

[3] P. Bhagwat, B. Raman, and D. Sanghi. Turning 802.11 inside-out. *ACM SIGCOMM Computer Communication Review*, Jan 2004.

[4] S. Buchegger and J. Y. L. Boudec. Performance analysis of the confidant protocol: Cooperation of nodes - fairness in dynamic ad-hoc networks. In *Proceedings of IEEE/ACM Symposium on Mobile Ad Hoc Networking and Computing (MobiHOC), Lausanne*, June 2002.

[5] W. Cheswic and W. Bellovin. *Firewalls and Internet Security*. Addison Wesley, 1999.

[6] Eric Cole. *Hackers Beware*. New Riding Publishing, 2001.

[7] D. Denning. An intrusion detection model. In *IEEE Transactions on Software Engineering*, volume SE-13, pages 222–232, 1987.

[8] S. Dharmapurikar, P. Krishnamurthy, T. Sproull, and J. Lockwood. Deep packet inspection using parallel bloom filters. In *Micro, IEEE*, volume 24, pages 52–61, Feb. 2004.

[9] M. R. Garey and D. S. Johnson. *Computers and Intractability*. W. H. Freeman and Company, 2000.

[10] S. Garfinkel and G. Spafford. *Practical UNIX and Internet Security*. O'Reilly and Associates, 2nd edition, 1996.

[11] C. Ko, P. Brutch, J. Rowe, G. Tsafnat, and K. Levitt. System health and intrusion monitoring using a hierarchy of constraints. In *4th International Symposium, Recent Advances in Intrusion Detection*, pages 190–204, Oct. 2001.

[12] H. Luo, R. Ramjee, P. Sinha, L. Li, and S. Lu. Ucan: A unified cellular and ad-hoc network architecture. In *Mobicom*, 2003.

[13] S. Marti, T. J. Giuli, K. Lai, and M. Baker. Mitigating routing misbehavior in mobile ad hoc networks. In *Mobile Computing and Networking*, 2000.

[14] John McHugh. Intrusion and intrusion detection. *International Journal of Information Security*, 2001.

[15] P. Michiardi and R. Molva. Core: A collaborative reputation mechanism to enforce node cooperation in mobile ad hoc networks. *IFIP-Communication and Multimedia Security Conference*, 2002.

[16] David Wagner Nikita Borisov, Ian Goldberg. Intercepting mobile communications: The insecurity of 802.11. In *Proceedings of MobiCom'01*, 2001.

[17] H. V. Poor. *An Introduction to Signal Detection and Estimation*. Springer-Verlag, 2nd edition, 1994.

[18] T. Ptacek and T. Newsham. Insertion, evasion, and denial of service: Eluding network intrusion detection. Technical report, an SNI Technical Report, http://www.aciri.org/vern/Ptacek-Newsham-Evasion-98.ps, Jan. 1998.

[19] R. Rao and G. Kesidis. Detecting malicious packet dropping using statistically regular traffic patterns in multihop wireless networks that are not bandwidth limited. In *Proc. IEEE Globecom, San Francisco*, Dec. 2003.

[20] R. Stewart, Q. Xie, K. Morneault, C. Sharp, H. Schwarzbauer, T. Taylor, I. Rytina, M. Kalla, L. Zhang, and V. Paxson. Stream control transmission protocol. *RFC 2960*, Oct. 2000.

[21] D. Subhadrabandhu, F. Anjum, S. Kannan, and S. Sarkar. Domination and coverage guarantees through distributed computation. In *Proceedings of 43rd Annual Allerton Conference on Communication, Control, and Computing*, Champaign, IL, Sep. 2005.

[22] D. Subhadrabandhu, S. Sarkar, and F. Anjum. Efficacy of misuse detection in adhoc networks. In *Proceedings of IEEE SECON'04*, Oct. 2004.

[23] D. Subhadrabandhu, S. Sarkar, and F. Anjum. Misuse detection with imperfect defenders in adhoc networks. Technical report, University of Pennsylvania Technical Report, http://www.seas.upenn.edu/~swati/publication.htm, Apr. 2005.

[24] D. Subhadrabandhu, S. Sarkar, and F. Anjum. A framework for misuse detection in ad hoc networks - part i. In *IEEE Journal on Selected Areas in Communications Special Issue on Security in Wireless Ad Hoc Networks*, 2006.

[25] D. Subhadrabandhu, S. Sarkar, and F. Anjum. A framework for misuse detection in ad hoc networks - part ii. In *IEEE Journal on Selected Areas in Communications Special Issue on Security in Wireless Ad Hoc Networks*, 2006.

[26] V. Vazirani. *Approximation Algorithms*. Springer, 2001.