# Secure Neighborhood Discovery: A Fundamental Element for Mobile Ad Hoc Networking

Panos Papadimitratos[1], Marcin Poturalski[1], Patrick Schaller[2],
Pascal Lafourcade[2], David Basin[2], Srdjan Čapkun[2], Jean-Pierre Hubaux[1]
[1] EPFL, Switzerland, {panos.papadimitratos, marcin.poturalski, jean-pierre.hubaux}@epfl.ch
[2] ETHZ, Switzerland, {patrick.schaller, pascal.lafourcade, basin, capkuns}@inf.ethz.ch

***Abstract*—Pervasive computing systems will likely be deployed in the near future, with the proliferation of wireless devices and the emergence of ad hoc networking as key enablers. Coping with mobility and the volatility of wireless communications in such systems is critical. Neighborhood Discovery (ND), namely, the discovery of devices directly reachable for communication or in physical proximity, becomes a fundamental requirement and a building block for various applications. However, the very nature of wireless mobile networks makes it easy to abuse ND and thereby compromise the overlying protocols and applications. Thus, providing methods to mitigate this vulnerability and to secure ND is crucial. In this article, we focus on this problem and provide definitions of neighborhood types and ND protocol properties, as well as a broad classification of attacks. Our ND literature survey reveals that securing ND is indeed a difficult and largely open problem. Moreover, given the severity of the problem, we advocate the need to formally model neighborhood and to analyze ND schemes.**

## I. Introduction

Over the past decade, wireless, mobile communication technologies have matured and been widely adopted. The number of cellular phones now exceeds by far that of wired phones; millions of nomadic users routinely connect to Wireless Local Area Networks (WLANs); and wireless devices are commonplace in homes, factories, and hospitals. At the same time, the emerging mobile ad hoc and sensor networking paradigms usher in a new type of network: devices form multi-hop topologies in a self-organizing manner, relaying packets from other devices across multiple wireless links (hops), and essentially become the network.

Several applications are enabled already by these developments or are expected in the near future. Wireless sensor networks are deployed for environmental and building monitoring. Mobile ad hoc networks are used in disaster relief operations, with "rolled-in" base stations and portable radios, as well as in tactical operations with a multitude of vehicle-, aircraft-, or personnel-borne wireless devices. Static ad hoc or mesh networks are being formed by home computers with roof-top antennas. Low-mobility ad hoc networks will enable (often delay-tolerant) communication in urban environments; examples include networks of hand-held devices, wearable devices, and radio frequency identifiers (RFID).

The device mobility and the volatility of wireless communication, most often across radio frequency channels, result in connections that are frequently established and torn down without prior notification. The challenge here is to *discover neighbors* that, depending on the supported application, can be: (i) devices directly reachable for communication, i.e. *communication neighbors*, or (ii) devices in close proximity, i.e. *physical neighbors*. Typically it is assumed that if two nodes can communicate directly, they are within each-other's communication range, and vice-versa. Proximity and communication however are not always related.

Protocols for *Neighborhood Discovery (ND)* serve as fundamental building blocks in mobile wireless systems. Clearly, ND enables (multi-hop) communication, as it is essential for route discovery and data forwarding. ND can also support a wide range of system functionality: network access control, topology control, transmission scheduling, energy-efficient communication, as well as physical access control. Given the critical and multifaceted role of ND, its security and robustness must be ensured: ND protocols must identify as neighbors only those devices that actually are neighbors, even in hostile environments.

Securing ND is however a hard problem. The very nature of wireless environments and mobile computing applications makes it easy to abuse ND and thereby compromise systems for which ND is a building block. A striking example is that of defeating an *identify-friend-or-foe* system [1]. An attack that has become known as the *"MIG-in-the-middle"* was mounted in the late 1980s against the South African air defense; approaching aircrafts that appeared on their radars had to identify themselves by providing a rapid response to a challenge message. Angolan MIG airplanes received such South African challenge messages, *relayed* them (via the Angolan air defense) to South African airplanes flying over a different region at the same time, and obtained their responses. These responses were relayed in real-time back to the MIGs, which in turn transmitted them, successfully masquerading South African airplanes. Unobstructed, the MIGs bombed their targets.

Beyond this exotic context, ND vulnerabilities can be exploited in many existing systems, as shown in Sec. III. Equally important, numerous new attacks against ND protocols are likely to emerge in future wireless systems. In anticipation of such vulnerabilities, we examine, in Sec. II, different notions of neighborhood and the properties that must be satisfied by any ND protocol. Then, we classify attacks in Sec. III, and survey proposed solutions in Sec. IV. Our investigation shows that the security of ND remains a largely open problem, despite

the numerous existing proposals. We conclude by suggesting the formal analysis of ND protocols as the next step on a roadmap towards provably secure neighborhood discovery.

## II. Types of Neighborhood

Devices in existing wireless and upcoming mobile ad hoc networks are diverse in their characteristics and functionality. To introduce the problem at hand, we abstract away numerous details and consider system entities to be generic *nodes*. Each node has a unique identity, a processing unit, and a wireless transceiver.

Nodes communicate over the wireless medium, based on the state of the medium and the capabilities of their transceivers. We do not dwell on the transceiver characteristics, unless needed (in Sec. IV-C). In general, beyond technical characteristics of the receiver (such as their sensitivity), parameters and factors that determine the ability to communicate include: (i) the power of the transmitted signal, (ii) the distance between the transmitting and (intended) receiving nodes, (iii) the ratio of the received power over that of noise and interfering signals, and (iv) and impairments of the wireless medium (such as fading or scattering).

### A. Communication and Physical Neighborhood

First, we define the *communication neighborhood* of a node $U$ as the set $\mathcal{C}(U)$ of nodes able to send information directly to $U$. In other words, a node $V$ is a *communication neighbor* of $U$ if and only if $U$ is able to receive signals transmitted by node $V$. Equivalently, we denote the $V$-$U$ communication (neighborhood) by stating that the $(V, U)$ (wireless) *link* is *up*. Otherwise, we say that $(V, U)$ is *down*.

The above notion of neighborhood focuses on communication. But, intuitively, *neighborhood* suggests closeness. Typically, if $U$ can directly receive information from $V$, it can expect that $V$ is within its *nominal communication range* $r$. Such a notion of proximity is captured by our definition of a second type of neighborhood. We define the *physical neighborhood* of a node $U$ as the set $\mathcal{P}(U)$ of nodes within physical distance $r$ from $U$.

It is important to clarify that communication and physical neighborhood are *not* equivalent in general. On one hand, communication neighborhood does not imply physical neighborhood. For example, a node $V$ that increases its transmission power, perhaps by upgrading its transceiver, exceeds the expected communication range, and thus places itself in $\mathcal{C}(U)$ but not in $\mathcal{P}(U)$. On the other hand, physical neighborhood does not imply communication neighborhood. Consider, for example, $V \in \mathcal{P}(U)$ that *cannot* send information directly to $U$ because of an obstacle (e.g., a wall); clearly, $V \notin \mathcal{C}(U)$. The two types of neighborhood are equivalent only under the idealized (thus, not realistic) *unit disk* communication model, which considers $U$ and $V$ communication neighbors if and only if their (geometric) distance is below $r$.

Note that physical neighborhood is by definition symmetric. However, communication neighborhood, as defined, may be *asymmetric*. Even if $V \in \mathcal{C}(U)$, $U$ is not necessarily able to send information directly to $V$ and would therefore not belong to $\mathcal{C}(V)$. For communication neighborhood to be symmetric, both links $(V, U)$ and $(U, V)$ must simultaneously be *up*.

### B. Partial and Complete Neighborhood Discovery

*Neighborhood Discovery Protocols* attempt to determine the neighbors (communication or physical) of a given node. Therefore, their main requirement is correctness: to identify only nodes that are actual neighbors, that is, to prevent the attacker from tricking nodes into accepting non-neighbors as neighbors. Verifying that a given node is indeed a neighbor could be viewed as a stand-alone part of secure neighborhood discovery functionality; we term this as *verification*. For example, a node could obtain neighborhood information in an insecure manner, but then perform verification to achieve secure ND.

In practice, ND protocols are only *partial*, as they may fail to discover (and verify) all neighbors. This is because it is difficult to guarantee message delivery in wireless networks. In general, an attacker can jam communication and thereby prevent the discovery of one, many, or even all nodes that would be otherwise part of the neighborhood.

This problem can be avoided in restricted operating environments or where anti-jamming or other measures guarantee the delivery of messages. We call a ND protocol *complete* when it discovers all *honest* (or *correct*) neighbors, that is nodes that abide with the protocol functionality. We restrict ourself to honest, correctly functioning participants because *dishonest* or *faulty* nodes can always refrain from participating in the protocol execution.

### C. Neighborhood as a Building Block

Neighborhood discovery enables different types of system functionality, as the following examples illustrate.

*Physical Access Control:* Receiving a signal from an RFID tag with a tag reader can be used to authorize the access of the tag bearer to a building or authorize the entrance of a vehicle in a highway segment. Signal reception implies the tag is at most within a system-specific predefined distance (e.g., a few centimeters or couple of meters) from the tag reader. Physical access control systems leverage on the range-limited communication capabilities of their hardware (tags), aiming essentially at physical ND.

*Network Access Control:* In general, access to network resources is granted only to registered users or devices. Nonetheless, direct communication with a dedicated system entity can be an important access control criterion in mobile wireless systems. For example, nodes obtain connectivity with the Internet only when they are in range of a WLAN Access Point (AP) or a cellular system base station. Here, access control relies on communication ND.

*Routing:* In multi-hop wireless networks, all types of data communication and dissemination (one-to-one, one-to-(m)any, or broadcast) rely on the notion of neighborhood. The neighbors of each node are always the ones that receive and forward control traffic and data to and from the node, for example, for route discovery and communication with another

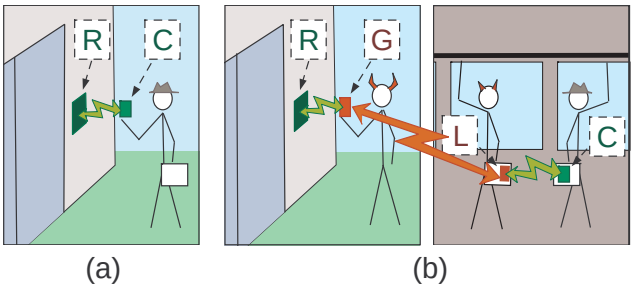


Fig. 1. RFID Access Control. (a) Normal operation: a legitimate user opens the door controlled by an RFID reader (R) using his RFID card (C). (b) The attack: the *leech* (L), next to the RFID card owner, and the *ghost* (G), next to the RFID reader, use a long-range link to relay transmissions between the card and the reader. As a result, the reader is misled to believe that the legitimate card is in its physical neighborhood and opens the door.

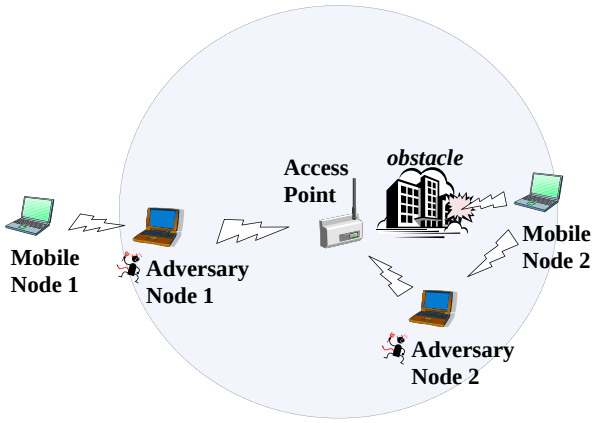


Fig. 2. Relay Attack on Network Access. Mobile Node 1 cannot directly connect to the Access Point because it is out of range. Similarly, Mobile Node 2 cannot directly connect to the AP due to an obstacle. However, the Attacker Nodes can act as relays between the Mobile Nodes and the Access Point, misleading them they can communicate directly, and thus control the communication.

(destination) node. If a destination is already identified as a neighbor, then no route discovery or calculation is necessary. In the case a position-based routing protocol is employed, the neighbor closest to the destination's position is selected. In all of these cases, communication ND is necessary. If efficiency or fault-tolerance are sought, a complete ND protocol would be desirable. For example, selecting the appropriate neighbor to forward data to or using alternative paths assume that many or even all neighbors have been discovered. Completeness of a ND protocol would therefore prevent the adversary from disconnecting nodes from their benign neighbors.

## III. ND Vulnerabilities and Attacks

Traditional security goals, such as authentication, secrecy, and non-repudiation, require reasoning about message contents. In contrast, secure neighborhood discovery relates primarily to the properties of the signals through which messages are exchanged. ND protocols operate essentially with respect to two layers: (i) an *abstract layer* that describes (benign and adversarial) message content handling, and (ii) a *physical layer* that describes (benign and adversarial) handling of signals sent across the communication medium (e.g., signal strength or time of arrival). Although physical and communication neighborhood are properties of the physical layer (distance covered by the signal or respectively the origin of the signal), the node's identity on the abstract-layer is used as the identifier in neighborhood relations. Note that this requires a binding between the "abstract" identity of a node and the transmitter (emitting the signal) associated to a node.

Attacks mounted against ND protocols can also be viewed in this layered manner. Consider a naive communication ND protocol commonly used when security is not a requirement: node $V$ sends a beacon "Hello, I am $V$". Upon receipt of this beacon, $U$ adds $V$ to its communication neighborhood. This protocol can be easily attacked on the abstract layer: an adversarial node $M$ can forge a message "Hello, I am $W$," convincing $U$ that $W$ is a neighbor even if this is not the case, thus violating ND correctness.

This naive protocol can be augmented with cryptographic mechanisms: $V$ can digitally sign the beacon with its private key, so that $U$ (or any other node that receives the beacon) adds $V$ to $\mathcal{C}(U)$ only when the signature is valid. However, $M$ could still receive a beacon from $W$ and simply replay it, to mislead $U$ into adding $W$ to $\mathcal{C}(U)$. Or, $M$ could achieve the same at the physical layer, relaying the signal carrying $W$s beacon on a per-bit or per-symbol basis. Although cryptography ensures that the received message (on the abstract layer) has been created by node $V$, the protocol does not guarantee anything about the physical layer of the communication, except that node $V$ must have emitted (somewhere, at some time prior to reception) a signal carrying the signed beacon.

Such *relay attacks* denote a fundamental way to attack ND protocols; we next discuss these, along with possible repercussions.

### A. Relay Attacks

Relay attacks, also known in the literature as *wormhole attacks*, are effective against the above-mentioned cryptographically augmented naive protocol. But they can also harm the more sophisticated ND protocols that we survey in Sec. IV. We discuss the implications of successful relay attacks next, in particular, for the upper-layer protocols and services discussed in Sec. II-C.

***Physical Access Control:*** We consider an attack against an RFID-based system that controls physical access to a building, illustrated in Fig. 1. For this attack [2], whose practical implementation was reported in [3], the adversary must control two nodes. The first adversarial node, the *leech*, is placed close to the victim's RFID tag. The second, the *ghost*, is placed next to the RFID reader controlling the building's

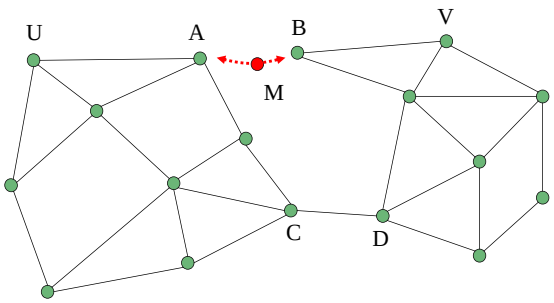


Fig. 3. Relay Attack on Routing (I). By creating an artificial link ($A$,$B$), the adversary $M$ attracts routes, e.g. $(U,V)$-routes that would otherwise use link ($C$,$D$). In this way, acting only locally, $M$ gains control over the communication of remote nodes, e.g. $U$ and $V$.

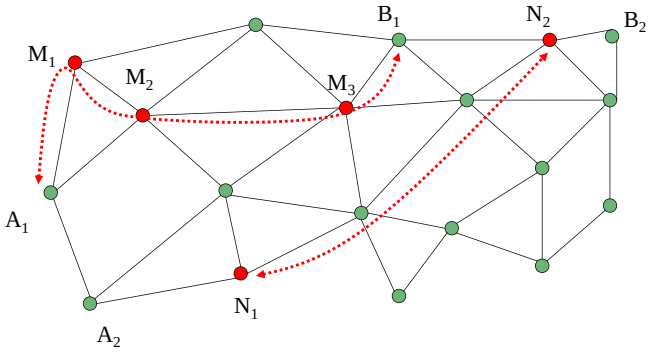


Fig. 4. Relay Attack on Routing (II). By relaying transmissions between nodes $A_1$ and $B_1$, the adversarial nodes $M_1$, $M_2$, and $M_3$ create an artificial, long-range link ($A_1$,$B_1$). Similarly, nodes $N_1$ and $N_2$ can use an out-of-band channel to relay transmissions between $A_2$ and $B_2$. In both cases, the artificial link offers a route much shorter then alternative ones, and thus attracts traffic the adversary has control over.

entrance. Independently of the distance from the victim's tag to the reader, the leech and the ghost relay messages between them, thereby misleading the reader into believing that the legitimate RFID tag is close, granting access to the ghost's bearer.

As pointed out in Sec. II, such a physical access control system should guarantee physical ND. However, it fails because its design considers communication and physical neighborhood as equivalent and it relies on a naive communication ND approach. As a result, the attack violates the correctness of the communication ND and thus the correctness of the physical ND.

***Network Access Control:*** We next consider an attack, illustrated in Fig. 2, against mobile nodes trying to connect to an Access Point (AP). As Mobile Node 1 (MN1) is out of the AP's range, Adversary Node 1 can easily act as a relay between MN1 and AP. A relay attack is also possible when two nodes are physical neighbors (with $r$ equal to the nominal communication range) but are not communication neighbors; this is the case for Mobile Node 2 and AP in Fig. 2. In both cases, the correctness of communication ND is violated.

One could argue that the adversarial nodes provide a service to the system, as they essentially extend the AP coverage. But, in doing so, the adversary takes control of the node-to-AP connections. It can then intercept the relayed messages, as well as modify and delete them at will. In wireless networks eavesdropping is easy, yet, without the adversarial relays there would be no communication to eavesdrop on. Moreover, data modification would be more difficult without the relay attack. If MN1 were in range of the AP, the adversary would need at least two strategically positioned and synchronized nodes: one node jamming the AP, to prevent it from receiving the messages of MN1, and the second node recording MN1's transmissions and replaying their modified version.

Eavesdropping can be prevented by encryption, whereas message modification and deletion can only be detected (for example, with the help of digital signatures and message sequence numbers) but not prevented. Nonetheless, the relaying adversaries can delete messages effectively and, more important, stealthily: unlike jamming, the victim nodes (notably, the sending one) can detect the message loss but not its cause. Even worse, the adversary can choose the point in time to delete messages in order to cause the most harm.

***Routing:*** Finally, we consider relay attacks against ND in a multihop ad hoc network, such as a sensor network. In Fig. 3, nodes $A$ and $B$ are close to each other but unable to communicate directly due to the terrain and their transceiver limitations. The adversary places a node $M$ within range of $A$ and $B$, where $M$ acts as a relay, making $A$ and $B$ believe that they are communication neighbors. Then, it is highly likely that $U$ and $V$ will communicate across a route that includes the adversary-controlled link $(A, B)$. Such a $U - V$ route would be shorter than one that includes $(C, D)$, and shorter routes are in general preferable. The result of this attack can be devastating. At first, the adversary-controlled link attracts considerable traffic. In addition, if the network relays a time-critical alarm, the adversary can stealthily cut-off its "link" and prevent the event detection by the network user.

The attacker's control over route establishment can be further enhanced, as shown in Fig. 4. $M_1$, $M_2$, and $M_3$ are nodes controlled by the attacker, acting as simple relays. At the same time, the adversarial nodes $N_1$ and $N_2$ relay messages across a private or off-line $N_1 \leftrightarrow N_2$ channel. Again, these attacks form short routes for many pairs of nodes, empowering the adversary to control significant amounts of network traffic.

### B. Classification of Attacks

To reason about the security properties of protocols, one must clearly define the attackers against which a protocol should achieve its goals. We classify possible attackers relevant to ND, as well as the solutions proposed in the literature, discussed in Sec. IV.

We begin by differentiating between *external* and *internal* attackers. In contrast to an external attacker, an internal attacker is an entity that is a legitimate participant of the network, typically possessing cryptographic keys as all honest participants do. With this distinction in hand, we can identify different types of attacks mounted by:

i) An attackers (internal or external) misleading two honest nodes that are not neighbors into establishing a neighborhood relation.
ii) An attacker (internal or external) tricking an honest node to believe that an adversarial node (internal) is its neighbor, although it is not.

The above attack types can have variants that involve a higher number of adversarial nodes.

An important characteristic of attackers is the delay they introduce when relaying messages. This is a critical parameter in the case where timing bounds are used in the defense against relay attacks, as explained in Sec. IV. We therefore classify attackers as:

i) *Slow relays*, if they need to receive the entire message before they are able to relay it.
ii) *Fast relays*, if they can start relaying a message only after receiving a portion of it.

From a practical point of view, a slow relay can be easy to implement, as an adversarial node that replays messages, operating at the abstract layer. In contrast, a fast relay, operating at the physical layer, would require sophisticated hardware, customized to relay the signals.

We also differentiate between *short-range* and *long-range* relay attacks. The former resulting in fictitious links shorter than the nominal node communication range $r$, and the latter in links longer than $r$. This distinction is meaningful because short-range relays, as opposed to long-range ones, do not violate the correctness of physical ND. As a result, they cannot be detected by mechanisms protecting physical ND.

We further differentiate relay attacks according to the adversarial node behavior: they can either always forward packets or do so selectively. Moreover, they can relay messages using omnidirectional or directional antennas. Finally, beyond relay attacks, the adversary can jam node communication in a selective or brute-force manner, possibly adjusting its transmission power and thus its impact. In the context of ND, jamming can obviously prevent the completeness of ND (even thwart the discovery of any neighbors at all), but also allow subtle attacks against some existing ND schemes (Sec. IV-B).

It is important to point out the difference between relay attacks and *tunneling* attacks [4], which were introduced in the context of routing. Suppose, for example, that two internal adversarial nodes participate correctly in their respective ND protocols, but "tunnel" (i.e., encapsulate and transmit to each other) control traffic, so that they appear as neighbors on routes discovered by the routing protocol. In contrast to relay attacks, tunneling attacks cannot be thwarted by any secure ND protocol.

## IV. Existing ND Approaches

We categorize schemes in the literature that could be used to secure ND according to their basic elements, analyze their properties in the light of our definitions and attack classification, and comment on their practicality.

### A. "Distance Bounding" Approaches

In addition to the use of cryptography, which establishes the identity of the node(s) participating in the ND protocol, Distance Bounding (DB) protocols estimate the distance to a potential neighbor $V$ by measuring the signal round-trip time and multiplying it by the signal propagation speed. For the case of radio frequency (RF) signals, which travel at the speed of light, it is essentially impossible for the (internal or external) adversary to decrease the estimated distance. As a result, DB protocols can guarantee *physical* ND: if the obtained distance is $r$, then the actual distance to $V$ is less than or equal to $r$.

If $r$ were the maximal communication range of a given transmitter and receiver pair, one might be tempted to declare that the DB approach can be used for communication ND. However, as argued in Sec. II, physical neighborhood does not always imply communication neighborhood. A short-range fast relay attack against two nodes that are physical but not communication neighbors, as in Fig. 3, would clearly violate the communication ND correctness.

On the contrary, if the adversary is only capable of slow relaying, i.e., it must receive an entire message before it can replay it, DB can detect the relay. Indeed, if the protocol uses sufficiently long messages, the relaying delay introduced by the attacker will result in an estimated distance exceeding $r$. With such a design, any of the DB protocols discussed below can be a candidate for *communication* ND. Even though DB protocols, as presented in the literature and discussed next, provide neighborhood verification, it is trivial to have (any non-secure) neighborhood discovery prior to the verification.

***Challenge-Response Delay Measurement:*** Two nodes can perform a *ranging* operation, that is, exchange messages along with their own measurements of the involved (processing and transmission) delays, and estimate their distance. Nonetheless, (authenticated) ranging cannot prevent a dishonest internal participant from falsifying its own measurements and thus violating the correctness of ND.

This vulnerability can be thwarted by a DB operation, first introduced by [5] and used later in numerous works, including its first use in mobile ad hoc networks by [6]. Brands and Chaum [5] propose to measure delays during a *rapid bit-exchange* phase, with nodes performing low-complexity calculations involving negligible processing delay. This implies that an internal adversarial node executing the protocol with an honest node cannot respond noticeably faster than honest

nodes and thus cannot convince its honest peer it is closer than it actually is.

Two or more colluding internal adversarial nodes can defeat most schemes descending from [5]: for example, an attacker neighbor to the victim could perform the rapid bit exchange but provide messages that corroborate the identity of another remote attacker. There are exceptions, such at the *Distance Bounding Proof of Knowledge* scheme found in [7], which guarantees ND under the assumption that colluding adversarial nodes do not share their secret keys.

***Message Time-Stamping**:* Under the assumption that nodes have precisely synchronized clocks, DB could be performed by estimating the delay for a one-way message exchange. This can be done with the addition of a time-stamp (and authentication of the messages) [8]. The scheme can be practical and effective only if the time-stamp value is set at the appropriate instant for an outgoing message. This would allow the receiver to measure the signal propagation delay plus some fixed (known) processing and transmission delays. Unlike challenge-response schemes, this protocol is secure only against external adversaries; internal adversarial nodes can forge arbitrary time-stamps.

We note that the practicality of distance bounding schemes has not been verified at the time this article is written. There is no proof-of-concept implementation for wireless networks. The implementation of a solution that is secure only against external adversaries appears easier, whereas an implementation that involves a rapid bit exchange poses additional challenges.

### B. Location-based Approaches

If available, trustworthy location information can be utilized for secure ND. If nodes executing a ND protocol are expected to provide their own location information, then it is straightforward to show that the related ND protocol would be at most resistant to external adversaries. This is the case for the two proposals surveyed next.

***Geographic Packet Leashes**:* Location-aware nodes can augment messages with an authenticated time-stamp and their location at the time of transmission [8]. With loosely synchronized clocks and knowledge of the maximum node velocity, the nodes can check if a received message originates from a sender not further than a given distance (e.g., the communication range). This guarantees physical ND. For communication ND, nodes could be equipped with a radio propagation model to determine whether or not a node at a given location is a communication neighbor. This requirement is impractical and hard to satisfy in general in a communication environment that is not known a priori and highly dynamic.

***Guard-based Wormhole Defense**:* If only a subset of nodes is location-aware, these nodes, termed as guards, can help other nodes establish neighbor relations [9]. Guards broadcast beacons reporting their location. Afterwards, other nodes exchange information about received beacons and assume they are neighbors if sufficiently many common beacons (at least some threshold $k$) were received. Relay attacks are detected based on two principles: (1) any beacon should be received at most once and (2) all locations in received beacons should lie in a circle with a radius two times the guard range. This can prevent relatively simple relay attacks. However, this is not the case for more elaborate attackers. For example, a selective wormhole can avoid detection based on principle (1). Moreover, one end of a wormhole can jam and prevent reception of legitimate beacons, relay beacons from the other end of the wormhole, and essentially "relocate" the victim node(s). The scheme is probabilistic in nature, and the threshold $k$ is calculated in the unit disk model, based on the density of guard and node deployment, resulting in an approximate physical ND.

### C. Directional Antennas Approach

The use of directional antennas for detection of relay attacks is proposed in [10], under the assumption of the unit disk model, the availability of antennas with an even number $n$ of non-overlapping zones each spanning an angle of $\frac{2\pi}{n}$, and the ability to have zones identically oriented for all nodes (e.g., using a compass). If two nodes are indeed neighbors, a message sent over some zone $z_i$ should be received at the opposite zone $\overline{z_i}$. Information (cryptographically protected) on the used zone is included in messages to detect relays in some cases. For increased protection, information can be exchanged among multiple nodes. This would ensure physical ND against at most two external adversaries. To the best of our knowledge, no proof-of-concept implementation of this scheme exists. In addition, its applicability is limited, as devices in many typical mobile computing scenarios use omni-directional antennas.

### D. RF Fingerprinting

Another approach relies on so-called Radio Frequency Fingerprinting, that is, the identification of characteristic signal patterns induced by radio transmitters. This allows the node receiving a signal to uniquely identify the source of that signal, that is, fingerprinting enables direct signal-origin authentication. Hence, under the assumption that the signal pattern is unforgeable, this guarantees signal reception directly from the claimed source and thus communication neighborhood verification. This approach is promising, but available results do not guarantee correct authentication with a probability close to one. Moreover, the infeasibility of forging RF fingerprints has not yet been shown. [11]

### E. Connectivity Approach

In multi-hop networks, local network connectivity information is proposed as the basis of a heuristic to detect wormholes and reject false links [12] and thus protect ND against external adversaries. The strength of the scheme is its practicality, in the sense that it does not require any specialized node hardware or capabilities. Nodes exchange locally communication neighborhood information, obtained through a non-secure ND mechanism. Afterwards they check for *forbidden structures*, that is, connectivity subgraphs that would exist if a wormhole were present (and would be unlikely otherwise). Forbidden structures depend on node density and

the connectivity model. Unless the density is low, simulation results show a 100% detection rate with no false alarms, for all connectivity models considered in [12] (unit disk, as well as more realistic models). However, the simulations assume a relatively naive relay, whereas a selective wormhole establishing only one or few fake links would be less likely to create a forbidden structure. Furthermore, although the wormhole detection scheme is evaluated, it is unclear how the ND scheme would perform. [12] points out that it might reject valid links.

### F. Centralized Approaches

There are related centralized approaches that do not attempt to provide secure ND, but rather try to detect the presence and approximate location of long-range wormhole and tunneling attacks. Such schemes rely on statistical [13] or visual analysis of the connectivity graph [14].

## V. Towards Provably Secure Neighborhood Discovery

Neighborhood discovery (ND) is central to many networking problems, especially in wireless networks with frequently changing topologies, and it enables a spectrum of new applications, such as authentication based on physical presence. The security requirements here are fundamentally different than those arising in classical entity or message origin authentication. The reason is that authentication pertains to the source of a received signal rather than the origin of a message. This leaves ample space for attacks against ND protocols that can have a devastating impact, leading to the compromise of the wireless mobile system functionality.

Our literature survey reveals that most of the proposed schemes provide physical ND but not communication ND. Protocols aiming at communication ND, which are based on physical ND protocols, often fail to achieve their objective. This is because these two types of discovery are *not* equivalent. At the same time, protocols for communication ND do not fully address the problem at hand. They are effective only under very specific operational conditions or they do not ensure correctness in all cases.

The complexity of the problem stems from the protocols themselves, as well as the need to clearly state the intended protocol goals and characterize the operating environment and intruder capabilities. As experience with Internet protocols has shown, these are highly nontrivial tasks. Many protocols once believed to be secure have been found flawed when formally modeled and analyzed. In the past decade, remarkable advances have been made in the automated analysis of standard security protocols, e.g., for authentication and key exchange [15]. The time is ripe to extend these formal models and methods, or develop new ones, and to analyze and secure ND protocols.

## References


[1] Ross J. Anderson. Security engineering: A guide to building dependable distributed systems. New York, NY, USA, 2001. John Wiley & Sons, Inc.

[2] Ziv Kfir and Avishai Wool. Picking virtual pockets using relay attacks on contactless smartcard. In *International Conference on Security and Privacy for Emerging Areas in Communications Networks (SecureComm)*, 2005.

[3] Gerhard P. Hancke. A practical relay attack on iso 14443 proximity cards. February 2005.

[4] Panagiotis Papadimitratos and Zygmunt J. Haas. Secure routing for mobile ad hoc networks. In *SCS Communication Networks and Distributed Systems Modeling and Simulation Conference (CNDS 2002)*, 2002.

[5] Stefan Brands and David Chaum. Distance-bounding protocols. In *EUROCRYPT '93: Workshop on the theory and application of cryptographic techniques on Advances in cryptology*, 1993.

[6] Srdjan Čapkun, Levente Buttyán, and Jean-Pierre Hubaux. Sector: secure tracking of node encounters in multi-hop wireless networks. In *ACM workshop on Security of ad hoc and sensor networks*, pages 21–32, New York, NY, USA, 2003. ACM Press.

[7] Laurent Bussard. *Trust establishment protocols for communicating devices*. PhD thesis, Thesis, Oct 2004.

[8] Yih-Chun Hu, Adrian Perrig, and David B. Johnson. Packet leashes: A defense against wormhole attacks in wireless ad hoc networks. In *IEEE Conference on Computer Communications INFOCOM*, 2003.

[9] Radha Poovendran and Loukas Lazos. A graph theoretic framework for preventing the wormhole attack in wireless ad hoc networks. volume 13, pages 27–59, Hingham, MA, USA, 2007.

[10] Lingxuan Hu and David Evans. Using directional antennas to prevent wormhole attacks. In *Symposium on Network and Distributed Systems Security (NDSS)*, 2004.

[11] Kasper Bonne Rasmussen and Srdjan Čapkun. Implications of radio fingerprinting on the security of sensor networks. In *International Conference on Security and Privacy for Emerging Areas in Communications Networks (SecureComm)*, 2007.

[12] Ritesh Maheshwari, Jie Gao, and Samir R. Das. Detecting wormhole attacks in wireless networks using connectivity information. In *IEEE Conference on Computer Communications INFOCOM*, 2007.

[13] Levente Buttyán, László Dóra, and István Vajda. Statistical wormhole detection in sensor networks. In Refik Molva, Gene Tsudik, and Dirk Westhoff, editors, *ESAS*, volume 3813 of *Lecture Notes in Computer Science*, pages 128–141. Springer, 2005.

[14] Weichao Wang and Bharat Bhargava. Visualization of wormholes in sensor networks. In *WiSe '04: Proceedings of the 2004 ACM workshop on Wireless security*, pages 51–60, New York, NY, USA, 2004. ACM Press.

[15] Alessandro Armando, et. al. The AVISPA tool for the automated validation of internet security protocols and applications. In *Proceedings of CAV'2005*, LNCS 3576, pages 281–285. Springer-Verlag, 2005.