From an IP Address to a Street Address: Using Wireless Signals to Locate a Target

Craig A. Shue[†], Nathanael Paul[‡], and Curtis R. Taylor[†]
[†]Worcester Polytechnic Institute, {cshue, crtaylor}@cs.wpi.edu
[‡]University of Tennessee and Oak Ridge National Laboratory, pauln@utk.edu

***Abstract*—How quickly can somebody convert an IP address of a target into a real-word street address? Law enforcement regularly has need to determine a suspect's exact location when investigating crimes on the Internet. They first use geolocation software and databases to determine the suspect's rough location. Recent research has been able to scope a targeted IP address to within a 690m (0.43 mile) radius circle, which is enough to determine the relevant law enforcement department that has jurisdiction. Unfortunately, investigators face a "last half mile" problem: their only mechanism to determine the exact address of the suspect is to subpoena the suspect's Internet Service Provider, a process that can take weeks. Instead, law enforcement would rather locate the suspect within the hour with the hope of catching the suspect while the crime is still on-going, which leads to stronger evidence and straightforward prosecution.**

Given these time constraints, we investigate how quickly an adversary can locate a target without any special law enforcement powers. Instead, we leverage the use of ubiquitous wireless networks and a mobile physical observer that performs wireless monitoring (akin to "wardriving," which seeks to search for wireless networks). We develop an approach that allows an adversary to send traffic to the target's address that can be detected by the observer, even if wireless encryption is in use.

We evaluated the approach in two common real-world settings. In one of these, a residential neighborhood, we used a single-blind trial in which an observer located a target network to within three houses in less than 40 minutes (with potential for more exact results using hardware such as directional antennas). This approach had only a 0.38% false positive rate, despite 24,000 observed unrelated packets and many unrelated networks. These results show significant promise for the geolocation strategy and demonstrate that adversaries with multiple potential observation points, such as law enforcement personnel, could quickly locate a target.

I. Introduction

It is relatively easy to determine a targeted Internet user's rough location using the target's IP address. Directed advertising often leverages the IP address of the user's machine and a geolocation database to tailor their marketing. On-going research has focused on increasing the precision of this geolocation. Recent work by Wang *et al.* [1] introduced a technique that can geolocate IP addresses with a median error distance of 690 meters, improving previous results.

It is far more challenging to determine a user's exact street address given an IP address without information from the Internet Service Provider (ISP). Law enforcement can subpoena information from the ISP, but it is often a slow process. Without ISP subpoenas, even the best geolocation work to date is insufficient at narrowing down the suspect pool. As an example, if we consider a populated portion of the United States, with roughly $57,000$ people per mi^2 according to the US Census [2], a 690m radius circle would encompass around 33,000 people. There are often practical and judicial constraints on the number of suspects that can be investigated for a single crime.

In this work, we ask: *How quickly can one convert an IP address into a physical street address? Can this be done without ISP support or law enforcement powers?*

To investigate these research questions, we use wireless networking to aid geolocation. Users are increasingly adopting wireless Wireless networking technology with recent market research studies estimating between 61% to 80% of US homes use wireless networks [3], [4]. Accordingly, if an adversary can use the wireless network to identify a target, then the adversary may be able to simply physically traverse the search area looking for the target. This approach is similar to the "wardriving" concept of mapping networks, but rather than mapping all networks, our approach is looking for a particular one.

Two factors make this exploration particularly interesting: wireless networks often use encryption and consumer-grade wireless routers often use features (e.g., Network Address Translation (NAT)) that can prevent the delivery of unsolicited packets. Rather than simply sending an identifying message to the attacker and looking for that same message with our mobile wireless observer, we develop techniques that overcome these obstacles.

In this paper, we present a system, which we call *Marco Polo*, that combines covert Internet signals with wireless analysis in order to remotely identify a target's geophysical location. We first narrow the search area using prior geophysical search techniques [1] and then dispatch a mobile observer that physically traverses the search space while monitoring the wireless spectrum

for signs of the wireless signature. Upon detecting the signature, the mobile adversary can detect the boundaries of the wireless signal and then use directional antennas to triangulate the target's wireless system's exact location.

This covert signaling process essentially forces the target's wireless LAN to issue beacons for pinpointing a target's location, similar to the namesake children's game of Marco Polo. This approach is viable since the search space can be exhaustively explored. These techniques are effective, even if the client uses wireless encryption and NAT devices.

The following are the main contributions of this work:

- **Covert Wireless Signals for Geolocation:** We introduce an adversary that seeks to locate a target using two components: a signaler and a mobile observer. We define and justify the adversary's abilities (Section III) and describe methods to allow covert, flexible, and reliable signals for the observer to detect (Section IV).
- **Practical Applicability:** We show our approach works beyond a laboratory environment by applying it in two realistic scenarios: 1) a single-blind vehicle-based scan of a residential neighborhood, and 2) a walking scan of an apartment building exterior and interior (Section VI). Our vehicle-based scan allowed our observer to find the target in less than 40 minutes, localizing the target to a set of three houses. We discuss how additional techniques, such as directional antennas and wireless signal strength measurements, could help determine an exact location and may decrease the search time required. Our apartment scan showed that an observer could determine the correct building from public roadways and further localize the apartment by traversing interior hallways in the building.
- **Privacy Countermeasures:** We discuss countermeasures to preserve the target's privacy. These include obvious measures, such as using hardwired connections or proxy devices. However, we also discuss more subtle countermeasures that could preserve the target's privacy without compromising convenience (Section VII).

II. Related Work

Several directions of related research exist. The most prominent area, IP address geolocation, has received significant attention. Another area, covert channels, blends its signal into a legitimate communication. Unlike traditional covert channels, we seek to derive location via side-channel techniques rather than through data exfiltration. We first explore one closely related work in greater depth and then briefly explore each of these related areas.

Chen *et al.* [5] attempt to find a wireless host that is behind a NAT device. The authors assume that the adversary will have the ability to subpoena Internet service providers (ISPs) for a suspect's location and the ability to arbitrarily intercept and manipulate packets destined to the target to identify the targeted machine. The authors shape the packets to be of specific sizes representing binary digits (their example uses packet size 100 to represent a binary '0' and 400 to represent a binary '1'). They then use a multi-channel wireless sniffer and error correcting codes to determine their signal.

Both our approach and the previous work aim to identify the target system using a wireless packet sniffer. Similarly, we use a notion of packet lengths to aid in identification. However, in our work, we: 1) do not require subpoenas or other ISP cooperation, 2) do not require man-in-the-middle abilities and need only to establish a TCP connection with the target if the target uses NAT, 3) do not assume the adversary is within wireless range of the target (we build on recent work to find a target [1] and then precisely locate the target's location), 4) use variable sized packets to quickly and robustly confirm the target's signal, and 5) can use multiple types of covert signals. These relaxed assumptions make our approach immediately practical for adversaries without a subpoena or other ISP cooperation.

A. IP Geolocation

A variety of work has focused on trying to roughly geolocate IP addresses [6]–[9]. Each of these provide varying degrees of proximity for geographical locations. Unfortunately, these approaches have error distances that are too large to exhaustively explore. However, recent work by Wang *et al.* [1] provides street-level accuracy by leveraging the fact that businesses often run web servers locally and provide their local addresses publicly. The authors use these web servers as landmarks to estimate the geophysical location of a target IP address to within a median error of 690 meters.

Our approach extends the previous IP geolocation work by taking a rough location area and finding the exact location of the target IP address using active probing. Under certain situations, with directional antennas and triangulation, one can identify the building or room in which a target is located along with its physical MAC address. To our knowledge, this was not previously possible with other IP geolocation work.

B. Steganography and Watermarking

Steganography, the study of secret messages inside of benign messages, has previously been explored in digital communication [10]. We use steganography to create watermark signals to break privacy in locating a

target. Previous work has investigated hiding messages in covert channels in a variety of protocols [11]–[15]. The key approach is to communicate using slight variations in protocol implementation.

III. Defining the Adversary

In this work, we show an approach that enables an adversary to geophysically locate a machine using a targeted IP address. This adversary can be viewed as having two components: an Internet-based signaler and a physically mobile observer. While we separate these components in our description for clarity, they can be combined in practice without affecting the adversary's success. This adversary has the following three abilities:

(a) the ability to communicate via the Internet,
(b) the ability to roughly geolocate a target's IP address,
(c) and the ability to physically scan the wireless spectrum of the geolocated physical region.

The adversary will follow a specific sequence of steps in an attack. We depict this process in Fig. 1, which we will reference as we discuss the details of these steps. First, the adversary will remotely connect to a target. Second, the adversary will craft signal packets that create a unique signature in the wireless radio spectrum in the target's local area network that can be detected by a mobile observer. The adversary and mobile observer can be different or the same person. Last, this mobile observer can then use the covert embedded signal to locate the target.

In the remainder of this section, we describe practical ways for an adversary to obtain the required resources to carry out a stealthy geolocating algorithm, and we describe optimizations in Section IV.

A. Internet Communication with the Target

When a target is directly connected to the Internet through a wireless access point, communication with the target becomes trivial. A wireless access point will wirelessly transmit any packet sent to the target's IP address. Even if the target discards the traffic via firewalls or other mechanisms, the adversary will have succeeded in having the packet manifest in the wireless spectrum of the target network, which is sufficient to create a covert beacon signal (Fig. 1, Signaler step 1). While this works well for wireless access points, the approach is more challenging when a wireless router is involved.

In residential settings, users may configure a wireless router to provide connectivity to multiple machines. To do so, these wireless routers employ NAT. When an internal network user initiates a connection to a remote host, the NAT device creates a mapping associating the internal and remote network IP addresses and transport layer ports. When the remote host responds, the NAT device consults its mappings to send the packet to the correct internal host. However, if no mapping exists for an incoming packet, the NAT device cannot determine the appropriate internal host and instead drops the packet.

Since NAT will drop unsolicited network traffic from the adversary, the adversary must somehow lure the target into initiating the connection with the adversary. In practice, the adversary has several options available. Adversaries may advertise servers with attractive content to lure a target into establishing a connection, such as the FBI's use of honeypots advertising illicit content [16]. Adversaries may also use peer-to-peer applications. Because NAT hinders peer-to-peer applications, these applications often have built-in NAT traversal techniques to allow peers to connect. Accordingly, when both the adversary and target run such software, the adversary can either directly connect to the target or advertise itself to the target, causing the target to initiate the connection.

Once the adversary is connected to the target, it can begin introducing the intended signal. In doing so, it may need to keep the client connected to keep the NAT mappings valid. In other cases, the adversary may be able to keep the NAT mappings in place, even after the connection is closed. We discuss these issues in further detail in Section IV.

B. Roughly Geolocating a Target

Once a connection is made with the target, the adversary must narrow the geophysical search space (Fig. 1, Signaler step 2). While IP geolocation has been studied using many different strategies, recent work by Wang *et al.* [1] is especially promising. The authors used latency measurements and landmarks to help locate targeted hosts. They analyzed three different IP data sets including a PlanetLab, residential, and online mapping dataset. The lowest observable median error rate, 690 meters, was for a PlanetLab data set where the nodes are publicly reported by Universities. Larger median error rates were observed for their residential (2.25 km) and online maps dataset (2.11 km).

These landmark geolocation approaches improve with greater landmark density. Adversaries with more landmarks would have better localization. In the case of law enforcement contexts, for example, officers (acting as an adversary locating a target) may volunteer to provide landmark services through their home ISP to aid in searches. These landmarks, in addition to public landmarks, may allow for more robust latency measurements to localize the geographic search region.

C. Wirelessly Scanning a Physical Search Region

Our goal of wirelessly scanning for a physical target is achievable with off-the-shelf wireless components. Many modern wireless adapters allow programs to place

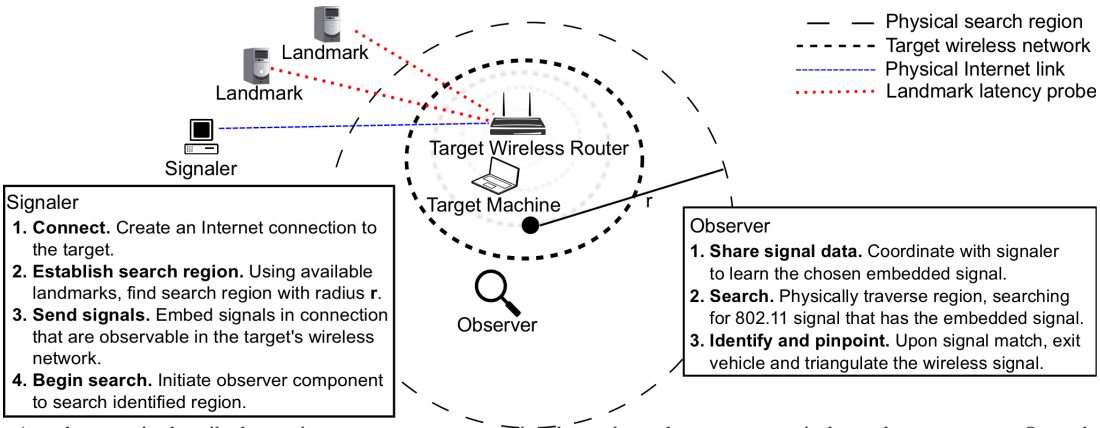


Fig. 1. An adversary is described as using two components, a signaler and an observer, to precisely geolocate a target. Once the signaler establishes a connection with the target, the signaler sends signal packets to the target via the connection. The observer will search a physical region (within radius r of the target) for wireless LAN activity that matches the signal signature.

the adapter into "monitor mode," in which the wireless adapter senses all radio traffic on a given wireless channel. If placed near a wireless signal emitter, this monitoring device can record all transmitted packets. Wireless security protocols, such as WEP or WPA, simply encrypt packets starting with the network layer. A monitoring device can still see the packet size and wireless MAC addresses of each communicated packet.

Digitally decoding multiple wireless signals in a reception range is a challenge (Fig. 1, Observer step 2). A typical commodity device can only tune to one wireless channel at a time. In North America, the 802.11 B and G protocols require 11 channels to be monitored. However, to accommodate 802.11 A and N in the 5GHz range, an additional 21 channels may need to be monitored. Chen *et al.* [5] addressed this problem by creating a multi-channel listening device by linking USB AirPcap dongles through a USB hub to a laptop. They were able to observe 11 channels simultaneously with excellent performance characteristics. This approach could be replicated with multiple units to simultaneously observe all the channels in use. An alternative, channel hopping, would also be viable. However, the speed at which the observer may search would be limited by the need to ensure each channel was observed long enough to detect the signal before continuing. We simply note the adversary has multiple viable options, depending on the available resources.

While the detection of wireless signals through mobile detection units is not a new concept, we face new challenges not found in previous work.

Both the FCC, with their use of Mobile Direction Finding (MDF) vehicles for locating pirate radio stations [17], [18] and the 2012 London Olympics, with a pedestrian enforcement team to locate unauthorized 802.11 wireless networks [19], have both demonstrated the practicality of detecting and triangulating radio signals.

While the concept is similar for our approach, our adversary's task differs in several important ways. First, the mere presence of a wireless networking signal is not itself sufficient evidence: while any pirate signal or unapproved wireless signal is inherently evidence of unauthorized transmission, most wireless transmissions that our adversary will encounter will ultimately be irrelevant, greatly increasing the challenge. Rather than finding the source of any detected signal, our adversary must find the source of a particular signal among many.

IV. Covert Communication Maintenance

The adversary's signaler component must be able to reliably transmit signal packets for the observer component to find. The adversary's goal is to embed a covert beacon signal into the connection to the target. When a target is directly connected to a network via a wireless access point, such signaling is trivial. However, with NAT in wireless routers, the adversary must somehow keep a NAT mapping fresh while sending the signal. A couple options are to 1) use in-band signaling by providing the target with content or using keep-alive messages or 2) use connection-less protocols that cause NAT boxes to maintain a connection with a remote machine and simply send packets regularly enough to keep the connections fresh.

While both of these strategies are viable, we introduce two novel techniques that are broadly applicable for maintaining NAT state (and thus eliciting wireless transmissions) while not affecting any user applications running on the target machine. We now describe these strategies in greater detail.

A. Out-of-Window TCP Signaling

For our initial approach of maintaining a connection with a target, we were inspired by the approach described by Handley *et al.* [20] that uses the time-to-live (TTL) field in the IP packet header to strategically drop packets

after traversing network middleware. In our case, we need a mechanism that drops a packet after it is transmitted wirelessly to the host, but before it reaches the user's application. Such a mechanism would allow application-agnostic signaling with less likelihood of detection.

The TCP protocol's sliding window implementation, which is designed to reorder and acknowledges packets, provides a suitable mechanism for strategic packet dropping. In particular, the targeted machine's TCP implementation will either silently discard or send duplicate acknowledgements to packets with out-of-window sequence numbers. In our testing, we found that Windows XP, Mac OS X, and Linux each dropped out-of-order packets without impact to network applications, such as telnet or HTTP. We could only detect these packets using packet capture software.

While out-of-window signal packets are discarded by multiple popular operating systems, we must confirm such packets will not be discarded by the target's router. We tested five routers, including a router from each of the four leading consumer-grade wireless router manufacturers, as shown in the second column of Table I. We found that out-of-window packets were forwarded wirelessly by all of the tested routers. While some other routers may detect and discard out-of-order packets, out-of-band signaling is practical for several popular consumer routers.

TABLE I
THE FIVE COMMODITY ROUTERS WE TESTED FORWARD OUT-OF-WINDOW TCP PACKETS AND ONLY ONE PREVENTED SIGNALING AFTER THE TCP CONNECTION TERMINATED.

Router Model	Forwards Out-of-Window Packets	Forwards after Termination
Belkin F5D8235-4	yes	yes
D-Link DIR-655	yes	yes
Linksys E900	yes	no
Linksys WRT54G	yes	yes
Netgear WNDR3700	yes	yes

B. Signaling after Connection Termination

All NAT devices must determine when to expire dynamic mapping. RFC 2663 specifically warns implementers to not simply delete a mapping when a TCP FIN or RST packet is seen, since there could be retransmissions. Instead, it recommends deleting a mapping after 24 hours of non-use [21]. RFC 5382 clarifies by recommending an idleness time-out no shorter than 2 hours. However, if either party sends a FIN packet, it states the mapping can be deleted only after 4 minutes of idleness [22]. Cisco uses a 24-hour expiration by default in their commercial-grade NAT devices and renews a mapping whenever it is used, regardless of FIN packets or whether the entry renewal comes from inside the network or outside the network [23].

An adversary can exploit these standards and guidelines to allow indefinite covert signaling, provided it sends a signal at least once every 4 minutes. This would allow an adversary to temporarily establish a connection with a target and then arbitrarily send signals as long as it desires, regardless of connection termination.

To determine whether this attack works in practice, we again tested five consumer-grade devices. Using raw sockets, we used a signaler that sent packets even after a related TCP connection had terminated. In the third column of Table I, we show the results of our testing. We found that four routers continue to send packets to the destination after the connection is closed, while one does not. This demonstrates that signaling is indeed viable on a set of popular consumer-grade wireless routers.

V. SIGNALING USING VARIABLE PACKET LENGTHS

Other geolocation covert signaling approaches are possible. By varying packet lengths, we can geophysically locate a target (Fig. 1, Signaler step 3, Observer steps 1 and 2). The 802.11 protocol transmits the packet length and the MAC addresses of the sender and destination in an unencrypted header for each packet, followed by the encrypted payload (if encryption is used). By sending the MAC addresses and length without encryption, wireless devices can quickly discard unrelated packets without cryptographic operations, conserving battery and computational resources.

The exposed packet length field and MAC addresses allow the adversary's observer to easily detect the participants and packet sizes of wireless communication. If the adversary's signaler sends specifically sized messages to the target, the observer component can detect whether these same sized packets are received by a wireless participant in its observable wireless spectrum area. By sending variously sized packets, the adversary can compute a confidence in how unlikely it would be for the pattern to occur in normal traffic, and the adversary could choose a pattern that should rarely occur in benign traffic.

The work by Lin *et al.* [24] shows that many applications have specific packet sizes they favor. For example, Apache HTTP was found to use packets of size 1216 for over 95% of its traffic, BitTorrent used packets of size 377 bytes for over 80% of its traffic, and eMule, another peer-to-peer application, used three specific file sizes for almost 70% of its traffic. The consistent packet sizes for these applications can be easily distinguished from the irregular packet sizing technique we use. For our experiments, we collected $135,076$ data-carrying wireless packets in a residential neighborhood. The observed packet sizes were skewed to small packets and large packets, with packets ranging from 750 to 1500 bytes being relatively rare in our snapshots. Accordingly, we create packets with sizes selected uniformly at random from 750-1500 bytes to decrease the risk of unrelated

activity producing signal values. This allows us to tune detection to reduce false positives even when near busy wireless networks.

Packet size signaling has significant benefits. The detection of the signal is straightforward. The signaler and observer can use a shared database of packet lengths and synchronized clocks to detect packets that are signaled. Likewise, false positives are easy to detect since they tend to be isolated and are not part of a long sequence of signals. Likewise, a longer correct sequence of events can improve an adversary's confidence of having detected the target.

VI. Packet Size Signaling in Practice

We implemented and experimentally evaluated the packet size signaling approach to determine whether it is practical as a covert detection signal. We used two real-world mobile observer tests to demonstrate the feasibility of the approach. We used a laptop with an inexpensive external omni-directional wireless adapter for our mobile tests. We previously tested and confirmed the approach works using both enterprise-grade wireless access points and with consumer-grade wireless routers, in both 802.11g and 802.11n wireless networks in the 2.4GHz and 5GHz bands, with both WPA2 personal and WPA2 enterprise wireless security modes. In our mobile experiments, we used WPA2 personal and consumer-grade wireless routers; we did not join the wireless networks or attempt to break the encryption of any transmissions.

We configured the adversary's observer system to use Kismet v2011-03-R2, a wireless network packet capture tool, on the monitor system. The Kismet tool can be used to place the monitoring system's wireless adapter into "monitor mode," in which the system sees each of the packets wirelessly transmitted. The system logs all the network activity it sees in the standard `tcpdump` packet capture format. By default, the Kismet tool uses channel hopping to continuously move across wireless channels. We configured Kismet to focus on the specific wireless channel in use by the client to allow continuous monitoring of the client. This experiment is similar to using multiple adapters watching each channel of the spectrum in parallel and doing per-adapter analysis on each. Channel hopping schemes could also be used with a channel hop delay proportional to the signal delays.

To prepare for our experiments, we generated our signal database in advance. In this database, we created two million records, each of which contained the signal value (the length of the packet to be sent) and the exact time that the signal would be transmitted. We then copied this database to the observer and then disconnected the observer from all networks. To accommodate small differences in the system clock values at the signaler and the observer systems, we provide a three second tolerance for comparing packets against the signal value. Accordingly, the observer must check the length of each packet it detects to see if the packet length matches any of the possible valid signal values. If so, the observer records the time, packet length, and source and destination MAC addresses from the packet header. For each destination MAC address, we determined the percentage of the last n transmitted signal packets that were actually sent wirelessly to the destination.

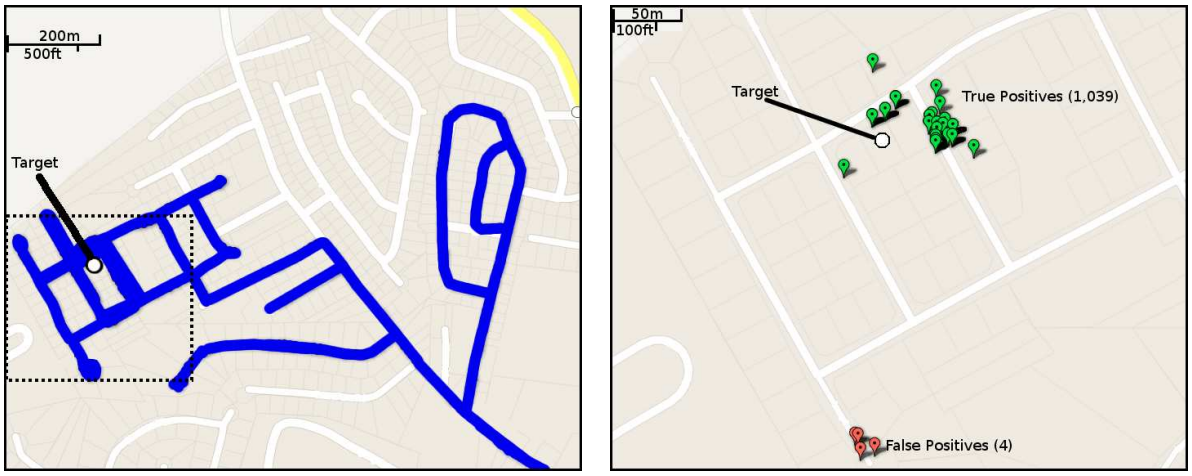
To show the practical applicability of the approach, we performed two types of experiments: 1) a single-blind driving geolocating test in a residential neighborhood, and 2) a walking test within and around a large apartment building. We again used commodity hardware for the wireless infrastructure and laptop systems for the search device. We configured the adversary to monitor the same wireless channel as the target, emulating an adversary observing each wireless channel simultaneously.

A. *Residential Neighborhood*

In our single-blind residential neighborhood experiment, one author, which we label the target, used an existing wireless network in a residential neighborhood. Another author, the observer, attempted to find the target. The observer was blind to where the target was located. The observer was only told the wireless channel the target would use (channel 11) and given a map of the rough area that the target was in (the map segment shown in Fig. 2(a)). Further, the observer knew the target would use a wireless link to connect to the adversary's signaler component via the Internet. The rough location area map represents the approximate error-bounds of current geolocation techniques.

The target used an existing wireless network in a residential neighborhood. To avoid biasing the experiment, the target did not reposition the router or laptop from their normal locations, despite clear obstacles that would hinder signal detection. The target's wireless router linked the target's laptop to a cable-based Internet connection. Through this wireless link, the target connected to a remote server that acted as an adversarial machine. The target left the connection active for the entire experiment.

To prepare for the experiment, the observer copied the timestamp and packet size database file to a local laptop. The observer then drove to the search area and configured his vehicle for observations. The observer's equipment included 1) a vehicle, 2) a laptop, 3) an external USB wireless adapter, which was taped to the car's passenger window, and 4) a GPS-enabled smartphone that recorded the vehicle movements with a clock synchronized with the observer's laptop. The observer then drove around the search area. Each time



(a) Search Space (traversed path shaded) (b) Signal Detection (dots indicate signal found)

Fig. 2. Using a single-blind search, our adversary was given the mapped area to traverse. Fig. 2(a) shows the attacker's search path through the area (shading indicates the path). Fig. 2(b), an enlargement of the dashed region in Fig. 2(a), shows the points where the target signal was detected. The map data and imagery are ©2012 Google. We digitally altered the maps to label the target location, highlight the adversary's path based on GPS coordinates, and to omit street names.

a packet matched the signal, the laptop would emit an audible alert, allowing the observer to focus on driving yet providing feedback to allow the observer to circle back and perform follow-up measurements in potentially interesting areas. The observer was only allowed to monitor from public roads and sidewalks to emulate practical usage.

In Fig. 2(a), we show the search space, using blue shading to show the path the observer took. We have labeled the target location for the reader's reference. In Fig. 2(b), an enlargement of the dashed region in Fig. 2(a), we show the locations where matches were found by the observer. The cluster near the bottom are false positives: the observer saw four packets that matched the signal from two unique computers, but did not see any subsequent matches in the area. The observer then continued searching and found a cluster of matches near the top. Each of these $1,039$ matches were true positives. The volume and consistency of the matches allowed the observer to have confidence that the observer found the signal. The source and destination MAC addresses detected by the observer were the correct MAC addresses for the target's router and laptop (the observer did not know these values in advance).

Before revealing the location of the target to the observer, the observer was asked about the target's likely location. The observer could not narrow it down to a specific house and instead indicated that a gap between two houses had the strongest reading. Through this gap, the rear portion of a third house was visible. The target was actually located in the rear portion of this third house. Physical obstructions (a masonry fireplace with metallic shielding and a large LCD panel television) were directly next to the wireless router, likely obstructing the signal in other directions from the house, hindering readings that would have further aided localization. In performing this experiment, we did not use directional antennas or signal strength meters, which may have allowed the adversary to determine the exact target location.

This experiment allows us to demonstrate the practicality of the approach. The observer was given a search space of roughly 1.23 km^2, while previous work has localized an attacker to 1.50km^2, causing our search area to be roughly 82% of realistic bounds. We were able to localize the likely target location to roughly 0.01 km^2 (about 3 houses) with our approach. The observer traversed roughly half of the search space, driving at around 10mph for 33 minutes before first finding a valid match. It then took the observer 4 additional minutes to find a location with a strong signal (allowing him to see 100% of packets in a 40 packet sliding window). During the experiment, the signaler sent about 9.6MB of signal, averaging about 4.38KB/s.

Importantly, during the experiment, the observer was able to exclude significant unrelated wireless network activity. At the target area alone, there were over 15 visible wireless networks, of which 4 were transmitting on the same channel as the target. The observer saw $24,030$ packets on the monitored channel that did not match the signal and could be ignored. Further, our false positive rate was only 0.38%. Given that 1,039 packets were a true positive, we can quickly eliminate these false positives in our search results. This experiment shows that an adversary can realistically examine a search space to localize a target. Further, multiple observers could operate in parallel to expedite localization, such as in the case of law enforcement usage.

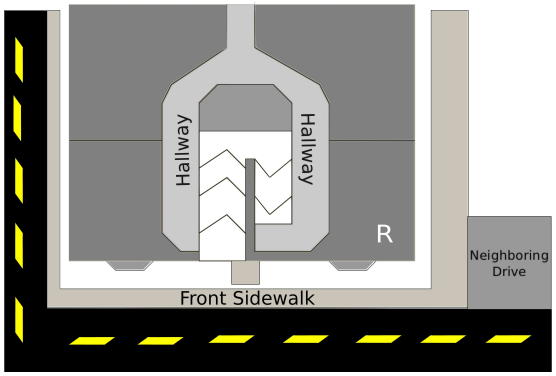


Fig. 3. Diagram of apartment building with floor layout and surrounding areas. The target's wireless router is labeled R.

B. An Apartment Building

In this experiment, a target was positioned on the second floor of a three-story apartment building, as shown in Fig. 3. The target was connected to an adversary signaler via its wireless network. An observer traversed the exterior of the building and interior hallways to determine the feasibility of an adversary trying to detect a target inside a multi-unit dwelling. All three levels of the apartment have the same design with four apartments per floor. The building has other similarly constructed buildings near by and there were approximately fifteen different wireless networks visible from within the building.

The observer was able to detect the target from various locations outside the building, including the neighboring public street. From outside the building, the wireless network was best detected along the front sidewalk but was also discoverable from the neighboring drive and roadway in front of the building which detected about 50% of beacon packets. The target was also located by running tests in the halls of all three floors. The detection rate was strongest on each floor in the quadrant nearest to the router, with true positive detection rates around 100%. However, the detection rate did not uniformly change with distance in many cases. The detection rate fluctuated greatly throughout the halls and had worse performance in hallways on the distant portion of the building. This is likely due to the construction of the building and signal deflection due to walls and stairwells. Despite these issues, and a few areas where the signal could not be detected, the observer was able to detect the target in the majority of the building.

This experiment demonstrates that the packet length method works well, even in an apartment environment. The successful discovery of the target from outside the building makes many use-cases feasible. Being able to see this type of apartment from a public roadway is a desired result because unauthorized entering of the building may not be possible in some cases.

VII. Countermeasures

We note that most Internet users are unlikely to employ countermeasures, due to lack of awareness of the risks or because they lack the necessary technical expertise. However, we nonetheless discuss countermeasures that could be employed by privacy-conscious individuals.

The most obvious approach to prevent detection is to use either a wired network or a proxy machine. The wired network will thwart our analysis since it eliminates the wireless signal. A proxy creates another level of indirection and the true target IP address may only be known to this proxy. While both countermeasures would be possible, they may be inconvenient for the target. Instead, we investigate countermeasures that could be integrated into current networking hardware transparently.

While modern wireless routers do not regularly fragment or resegment packets, the target could change its wireless router to enforce policies that alter all packet sizes. While enforcing constant-sized packets with padding is feasible, it could be detected easily. A router that uses variable-size packet reshaping would thwart the packet-length detection approach. However, it does not eliminate identification through more primitive approaches (such as sending data bursts resembling Morse-code). To protect against someone using data bursts while enforcing constant packet sizes, the target can saturate its bandwidth to hide a possible signal, but at a significant performance cost.

Anomaly detectors could also be used by a target to detect out-of-window packets and abnormal traffic shaping by its peers. Upon detecting an anomaly, the target could request its wireless router or AP to filter traffic from that source. However, an adversary with multiple machines may be able to continue the attack from a different source.

VIII. Conclusion

We have shown that one can convert an IP address into a physical location in less than an hour without any special law enforcement powers or support from ISPs. While this is a preliminary work, showing the potential of the approach, follow-on work could provide more exact location information by incorporating directional antennas and signal power meters.

Finally, we note the legality of this approach may vary by jurisdiction. Recently, a US federal judge ruled that unencrypted wireless communication is "readily available to the general public" and is thus legal to record under an exception of the Wiretap Act [25]. Since our approach also uses unencrypted headers, it may likewise be considered to fall under this exception. However, further case law and judicial rulings may be required to provide clarity as to whether the techniques we develop may be legally applied in any given jurisdiction.

REFERENCES

[1] Y. Wang, D. Burgener, M. Flores, A. Kuzmanovic, and C. Huang, "Towards street-level client-independent IP geolocation," in *USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, 2011.

[2] US Census, "Guide to state and local census geography," www.census.gov/geo/www/guidestloc/pdf/All_GSLCG.pdf, 2011.

[3] Business Wire, "Strategy analytics: A quarter of households worldwide now have wireless home networks," http://www.businesswire.com/portal/site/home/permalink/?ndmViewId=news_view&newsLang=en&newsId=20120404006331&div=-1063439563, April 2012.

[4] iGR, "Wifi bandwidth use in the U.S. home forecast to more than double in the next four years," http://www.marketwire.com/press-release/wifi-bandwidth-use-us-home-forecast-more-than-double-next-four-years-1589239.htm, November 2011.

[5] Y. Chen, Z. Liu, B. Liu, X. Fu, and W. Zhao, "Identifying mobiles hiding behind wireless routers," in *IEEE INFOCOM*, 2011, pp. 2651–2659.

[6] M. Freedman, M. Vutukuru, N. Feamster, and H. Balakrishnan, "Geographic locality of IP prefixes," in *ACM SIGCOMM Internet Measurement Conferences*. USENIX Association, 2005, pp. 13–13.

[7] B. Gueye, A. Ziviani, M. Crovella, and S. Fdida, "Constraint-based geolocation of Internet hosts," *IEEE/ACM Transactions on Networking*, vol. 14, no. 6, pp. 1219–1232, 2006.

[8] B. Wong, I. Stoyanov, and E. Sirer, "Octant: A comprehensive framework for the geolocalization of Internet hosts," in *Proceedings of the NSDI*, vol. 7, 2007.

[9] C. Guo, Y. Liu, W. Shen, H. Wang, Q. Yu, and Y. Zhang, "Mining the web and the Internet for accurate IP address geolocations," in *IEEE INFOCOM Mini-Conference*. IEEE, 2009, pp. 2841–2845.

[10] S. Katzenbeisser and F. Petitcolas, "Information hiding techniques for steganography and digital watermarking," in *Artech House Inc.*, 1999.

[11] J. Giffin, R. Greenstadt, P. Litwack, and R. Tibbetts, "Covert messaging through TCP timestamps," in *Privacy Enhancing Technologies*. Springer, 2003, pp. 189–193.

[12] G. Fisk, M. Fisk, C. Papadopoulos, and J. Neil, "Eliminating steganography in Internet traffic with active wardens," in *Information Hiding*. Springer, 2003, pp. 18–35.

[13] S. Murdoch and S. Lewis, "Embedding covert channels into tcp/ip," in *Information Hiding Workshop*, 2005.

[14] G. Shah, A. Molina, and M. Blaze, "Keyboards and covert channels," in *USENIX Security Symposium*, vol. 15, 2006.

[15] S. Zander, G. Armitage, and P. Branch, "A survey of covert channels and countermeasures in computer network protocols," in *IEEE Communications Surveys*, 2007.

[16] N. Anderson, ""the hidden side of your soul": How the FBI uses the web as a child porn honeypot," http://arstechnica.com/tech-policy/news/2012/04/the-hidden-side-of-your-soul-how-the-fbi-uses-the-web-as-a-child-porn-honeypot.ars, April 2012.

[17] M. Clyburn, "Blog of Commissioner Clyburn," http://reboot.fcc.gov/commissioners/clyburn/blog?entryId=932133, October 2010.

[18] G. Tristani, "Keeping the local in local radio," http://transition.fcc.gov/Speeches/Tristani/spgt811.html, September 1998.

[19] D. Cooper, "You've heard of the child catcher? meet the WiFi snatcher," http://www.engadget.com/2012/08/02/olympic-wifi-snatcher/?a_dgi=aolshare_twitter, Aug. 2012.

[20] M. Handley, V. Paxson, and C. Kreibich, "Network intrusion detection: Evasion, traffic normalization, and end-to-end protocol semantics," in *Proceedings of the 10th conference on USENIX Security Symposium-Volume 10*, 2001.

[21] P. Srisuresh and M. Holdrege, "Ip network address translator (nat) terminology and considerations," IETF RFC 2663, August 1999.

[22] S. Guha, K. Kiswas, B. Ford, S. Sivakumar, and P. Srisuresh, "Nat behavioral requirements for tcp," IETF RFC 5382, October 2008.

[23] J. Doyle and J. D. Carroll, *Routing TCP/IP*, ser. CCIE Professional Development. Cisco Press, April 2001, vol. 2, ch. 4.

[24] Y. Lin, C. Lu, Y. Lai, W. Peng, and P. Lin, "Application classification using packet size distribution and port association," *Journal of Network and Computer Applications*, vol. 32, no. 5, pp. 1023–1030, 2009.

[25] Dist. Court, ND Illinois, "In re innovatio IP ventures, LLC patent litigation," MDL Docket No. 2303, Aug. 2012.