On Location-Restricted Services

Eran Gabber and Avishai Wool
Bell Laboratories, Lucent Technologies Inc.

Abstract

Monitoring the location of user equipment is an important problem in many industries, including direct broadcasting satellites and others, where the physical location of the user determines the availability of the service or is crucial for the security or operation of the service. In this article we study four schemes for detecting the movement of user equipment, such as a set-top terminal, wireless local loop (fixed wireless) phones, and other "nonmovable" equipment, using existing (or emerging) communication infrastructures. The first two schemes are network-centric, which means that the network infrastructure determines the location. The two other schemes are terminal-centric, which means that they rely on the user's device. We start with the currently used scheme, which is based on the telephone network's caller ID features, and show how it can be undermined. Then we describe three more robust schemes: one that uses the cellular phone's enhanced 911 service, one that uses the Global Positioning System, and one that measures the time-difference-of-arrival of the satellite's broadcast. We discuss the accuracy, features, and vulnerabilities of each scheme. We also present possible attacks on these schemes that allow the attackers to conceal their movement, and evaluate the complexity and cost of the attacks.

Many services are inherently restricted to a certain geographic area due to various financial, copyright, and political issues. In order to enforce these restrictions, the providers of such services would like to detect the unauthorized movement of equipment which is supposed to remain stationary.

A good example is the satellite TV industry, where the content providers would like to limit the reception of the broadcast signal to certain geographic locations, and would like to prevent unauthorized movement of customers' receivers (also called *set-top terminals* or STTs) from a home to a public venue, or across an international border. Another example is preventing wireless local loop (fixed wireless) phones from being used as unauthorized mobile cellular phones. Detecting the movement of surveillance or monitoring equipment is another possible application. Other services may require a reliable indication of the customer's physical location, maybe as an added security measure against impersonation.

Security based on physical location should complement other security measures, such as encryption and authentication. It cannot replace other authentication schemes, since, as we shall describe, a determined attacker may "lie" about his or her location in various ways. However, once the cost of breaking a system is higher than the benefit of breaking it, the system has achieved its purpose.

In this article we describe four schemes by which a service provider may obtain the location of user equipment, using existing or upcoming technologies. For each scheme we provide an analysis of its accuracy and security, and suggest ways pirates may thwart it. None of our schemes is perfect; the pirates may spoof all of them using hardware tampering techniques such as those described in [1]. However, the schemes differ in the effort and cost required of the pirate.

In the rest of the article we will concentrate on the problem of locating STTs, since this is the most common use of location tracking today. The same techniques could be used for location tracking of other types of devices. We will also use the terms *customer* and *user* interchangeably.

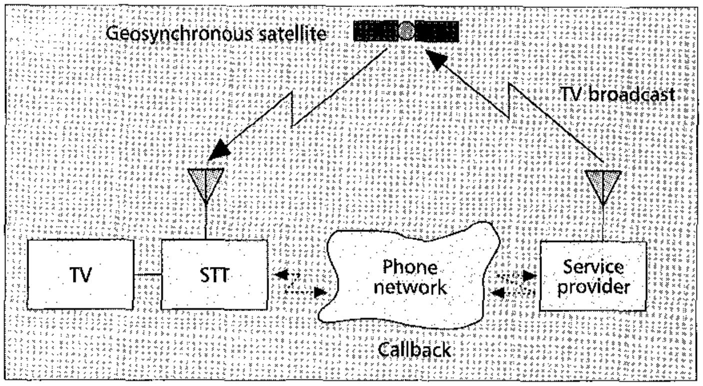
The first scheme we describe is the one currently used by Direct Broadcast Satellite (DBS) service providers. It is based on the caller ID feature of telephone exchanges (more precisely, it is based on the ANI or CND features). The three other schemes we describe use various techniques: the enhanced 911 (E911) feature of cellular networks, Global Positioning System (GPS), or time difference of arrival from the satellite signal. We believe that all of these schemes are more secure than the caller ID scheme. In all four schemes, the corresponding STTs are quite inexpensive to build and use commercially available components.

We do not consider schemes that rely on infrastructure which is not readily available. For example, we cannot expect a new GPS system that encodes additional provider-specific information in the satellite signal. Such a system would strengthen the GPS location scheme, but its cost would be prohibitive.

Tracking the movement of STTs may be misused in several ways. Customers may regard it as an invasion of their privacy,

0890-8044/99/$10.00 © 1999 IEEE

and as a form of "big brother" control over the movement of free citizens. Customers typically agree to report their exact location for receiving services that require connection to fixed infrastructure, such as electricity, water, gas, and wireline telephone. However, they usually are not required to report their location for receiving broadcast information, such as radio and television. Tracking information may also be misused by totalitarian governments that would like to control access to information by disabling foreign STTs imported to the country. This kind of "thought police" requires collaboration with foreign service providers. We do not address these privacy and censorship concerns in this article. We note, however, that customer privacy should be considered together with the commercial concerns of satellite TV providers. Note that these privacy issues are mostly specific to the satellite TV application, and are of less concern in the wireless local loop and other applications of location tracking.



■ Figure 1. *The components of a DBS system.*

The rest of the article is organized as follows. The following section describes several application areas of location tracking. We describe two network-centric schemes that rely on the network infrastructure to provide the location. We describe the caller-ID scheme currently used by the service providers, and discuss ways of undermining it. We present a more secure scheme based on the E911 feature of wireless cellular phone networks. Then we go on to describe two terminal-centric schemes. We show how to use the GPS for location detection, and discuss this scheme's vulnerabilities. We then present a scheme based on satellite ranging, which has many advantages over the previous schemes. We discuss the consequences of the pirates' ability to accurately clone an STT, and then conclude.

Applications of Location Tracking

This section describes several applications of location tracking. In all these applications, it is necessary to detect the unauthorized movement of equipment which is supposed to remain stationary. There is a large class of such applications, which include the DBS industry, wireless local loop (WLL) equipment, and surveillance equipment.

Direct Broadcasting Satellite Systems

As we already mentioned, the most widespread use of location tracking is in DBS systems. In such systems, the service provider transmits the TV programming contents to a geosynchronous satellite, which broadcasts it back to the customers. Since service providers of such systems derive their income from the customers' subscription fees, the broadcast is encrypted. DBS customers have an STT which receives the encrypted broadcast and decrypts the programs the customer is entitled to view. The decryption itself is performed by a secure, tamper-resistant module. Figure 1 depicts the components of a DBS system.

Piracy is a major problem for DBS service providers. For example, [2] claims that European service providers are losing £500 million a year in unpaid subscription fees due to pirated decoders. Therefore, the management of the decryption keys is central to the design of such systems [3, 4]. An important aspect of key management is how the keys for the next billing period are downloaded into the customer's STT. Modern DBS systems typically use a *callback* (or *return path*) scheme for this purpose: Once every billing period, the STT makes a phone call to the service provider, authenticates itself, and downloads the new keys.[1]

In many cases service providers would like to monitor the locations at which their customers install their STTs, and more important, to detect when customers move their STTs to new locations. The reason is that in some scenarios moving the STT may be a form of piracy.

The most common of these scenarios is when customers in country X cannot legitimately buy an STT for a service originating in country Y, even though the satellite's signal is received in country X. This may occur due to various financial, political, and copyright restrictions. Such situations typically lead to what [2] refers to as a "grey market," in which the STTs are bought in country Y and imported (or smuggled) into country X. Thus, the service provider would like to detect such activities in order to ensure that STTs moved to country X would not function there.

Moving STTs may be a form of piracy even without crossing international borders. A service provider would also like to restrict the movement of an STT from a customer's residence (where a subscription is cheap) to a commercial venue such as a theater or a bar.

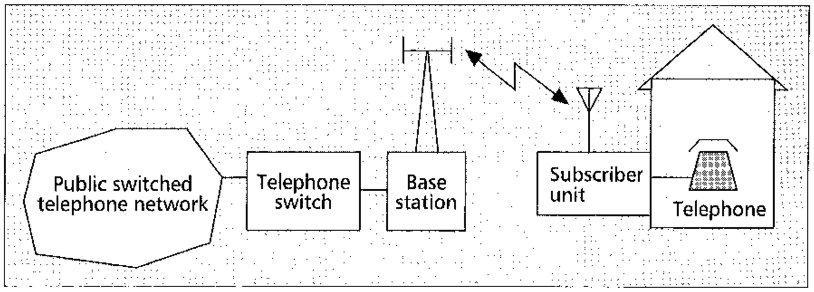
Wireless Local Loop Systems

WLL systems provide telephone services to customers by replacing the physical wire between the customer and the phone company's local office with wireless communication. WLL systems offer many benefits over traditional wireline systems. These include fast deployment, availability in rural and undeveloped areas, and no need for rights-of-way privileges for laying wires. WLL systems are especially attractive in underdeveloped countries [6], and for new phone companies that would like to provide phone service to customers without using the established local phone company's existing infrastructure. For a survey of current wireless local loop systems see [7].

Figure 2 describes the components of a WLL system. The subscriber has a receiver unit attached to the outside wall of his or her home. The unit connects to a regular telephone on one side; on the other side it communicates with the base station using some wireless technology.

Most equipment manufacturers use standard cellular telephone technology for WLL applications. Essentially, the subscriber's unit is a cellular phone with a modified interface, which is placed in a weatherproof box and attached to the customer's wall. Typically, the base stations in a WLL system are standard cellular base stations, and as such they even support customer mobility, despite the fact that the customer's units are supposed to be stationary.

[1] *This description is accurate for U.S.-based systems. STTs of existing European DBS systems, and the guidelines of the upcoming DVB standards, do not use callbacks [5].*



■ Figure 2. *The components of a wireless local loop system.*

In some countries, customers of WLL systems realized that their receiver units would still work after being moved, and be functional even inside moving vehicles. Therefore, some customers remove the WLL units from their homes and use them as improvised, somewhat bulky, mobile phones (e.g., car phones).

However, from the phone company's perspective, this behavior is a form of fraud. WLL customers who engage in this activity benefit from cellular phone mobility, but only pay the lower landline phone rates. Therefore, the phone company would like to detect unauthorized movement of WLL units so that appropriate action may be taken: for instance, the offending customers can be disconnected, or billed at a suitable rate.

Other Applications

Preventing tampering with surveillance and monitoring equipment, including unauthorized movement, has become crucial in recent years. For example, many arms control agreements are based on automatic surveillance equipment that is left at strategic locations (e.g., a video camera placed at an adversary's nuclear arms facility). The adversary can evade detection if the equipment can be disabled, or moved to a location where only approved activity takes place.

The same type of application is also relevant in various regulated industries. For instance, government regulations may require pollution sensors to be placed in the chimneys or sewage outlets of polluting industrial plants. Again, the plant may prefer to disable the sensors occasionally, for example, in order to burn cheaper but pollutant-rich fuel on a particular day.

Clearly, in these types of monitoring applications it is crucial to ensure that the sensing equipment is not tampered with. This requirement can be enforced using physical means (e.g., making the sensors difficult to detach) and administrative means (frequent inspections). Nevertheless, it is still necessary to ensure that the sensors have not been moved between inspections.

Other services may require a reliable indication of the customer's physical location, maybe as an added security measure against impersonation. For example, the customer may register her home address with service providers so that certain activities are enabled only when the customer is calling from her current home address.

Using the Caller ID Feature

This section describes the caller ID location scheme, which is currently used by many DBS service providers in North America for location tracking of STTs. The caller ID scheme is applicable only when the stationary equipment has a landline telephone connection, and the local phone company supports caller ID features.

How Caller ID Works

Modern telephone switches provide two features for identifying the calling party: the automatic number identification (ANI) [8] feature, which transfers the calling party's identification to another phone switch, and the calling number delivery (CND) [9] feature, which transfers the calling party's phone number to the callee. Usually, but not always, the ANI and CND contain the same information. The CND feature is commonly known as the caller ID feature. It has the property of being able to be enabled and disabled by the caller by dialing certain codes before the number. In contrast, the ANI feature is automatic and cannot be disabled by the caller. The ANI feature is currently used to locate callers to emergency services (911 in the United States) and for billing of calls to toll-free numbers. Since ANI is readily available for calls to toll-free numbers in the United States, many service providers currently use it to obtain the location of the caller. In the following discussion we use the popular name caller ID for either ANI or CND.

In the caller ID scheme, the service provider maintains a database of the customers' phone numbers. Whenever a customer calls the service provider, the service provider can verify that the call is coming from the designated phone number for this customer.

Some service providers, such as pay-per-view cable channels, even use caller ID to *identify* their customers; when a customer wishes to view a specific show, she must call from her home telephone (the number on the record) and punch a code that identifies the show. The service provider uses the caller ID to identify the customer and her STT without asking the customer for additional identification.

For the purpose of detecting STT movement in DBS systems, the typical solution is to piggyback a caller ID verification onto the callback. Recall that the STT needs to make a phone call to the service provider at least once every billing period in order to download keys and upload usage information. During this callback the STT and the service provider run a cryptographic protocol between them which authenticates both parties, guarantees the integrity of the data, and prohibits eavesdropping. Verifying that the STT has not moved is an easy addition to this protocol. As part of the authentication, the service provider matches the phone number obtained from the caller ID against the number on record for this particular customer's STT.

Caller ID is common in North America, less so in other parts of the world. We expect that the telephone systems in most nations will provide caller ID capabilities in the near future as those systems are upgraded.

Caller ID Accuracy

The caller ID feature is an indirect method of detecting the location of the STT, since it implicitly identifies a phone number with a geographic location. Therefore, the accuracy of the caller ID scheme really depends on the *inflexibility* of the local phone company.

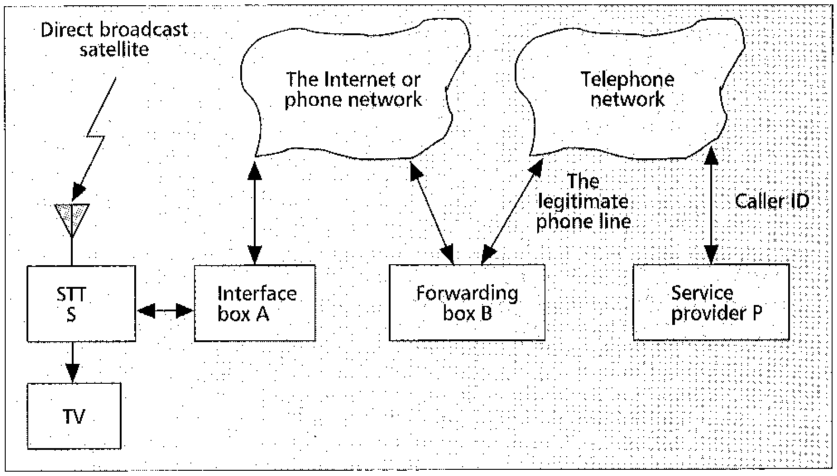
However, obtaining a location by caller ID is not very reliable. For instance, using a long-range cordless telephone would allow the STT to initiate the call up to half a mile away from the phone's base station, which is connected to the phone line at the legitimate location.

Caller ID Vulnerability

There are a few methods for moving the STT without the service provider's knowledge.

One of them is to relocate the phone line to a new address without changing the number (so-called number portability). Currently, local phone companies may be able to do this as long as the new address is served by the same phone exchange.

A more sophisticated attack, depicted in Fig. 3, can be used to move the equipment further. This scheme calls for an interface box, A, which is connected to STT S, and a forwarding box B, which is connected to the telephone line expected by service provider P. Box A is responsible for creating a communications tunnel to box B, which makes the actual call to P. This tunnel can be carried over the telephone network or over the Internet. In the first case, box B requires two phone lines: one to receive the call from A, and the other to call P. In the second case, both A and B can be implemented by PCs with modems that are connected to the Internet. To borrow from the Internet world, A and B implement the functionality of an HTTP proxy for an arbitrary protocol carried over a phone line.



■ Figure 3. *Defeating the caller ID scheme.*

The system operates as follows: A intercepts the call from S, captures the phone number of P, forwards this number to B, while presenting a ringing signal to S. B receives the number from A and makes the call to P. Once P answers the call, B notifies A, which notifies S. From now on, A and B pass the information verbatim between S and P. Note that no modification to S is needed, and no knowledge of the protocol between S and P is necessary to use this scheme.

We can make do with only one phone line for B if the communication protocol between S and P is simplex (only one side is talking at a time). In this case, A calls B, passes P's phone number, and immediately hangs up. Then B calls P. After a short period of time, A calls B again. This time B uses the call waiting feature to disconnect from P and connect to A. B should alternate between the two calls according to the expected origin of the next message, and pass the message to the appropriate party.

The cost of building A and B is in the order of a few tens of dollars. In fact, we have recently learned that cheap call forwarding devices, similar to our box B, are available off the shelf [10]. If one is using PCs connected via the Internet, which are also equipped with telephone modems, there is no extra hardware needed.

Another attack is possible when the STT is connected to a private branch exchange (PBX) which is connected to the telephone network. In this configuration the PBX is responsible for generating ANI and CND, which are passed to the callee. The pirate may purchase a PC-based PBX, and instruct it to generate the ANI and CND corresponding to the legitimate location, hiding the fact that the STT is actually connected to a different telephone line. This attack is not cost-effective, since PBXs tend to cost much more than STTs.

Note that call forwarding services offered by phone companies typically cannot be used to disguise the origin of the call, since the receiver obtains the caller ID of the ultimate originator of the call.

Network-centric schemes, such as caller ID or cellular E911 schemes (to be discussed later), are also vulnerable to attacks against the network itself. A determined hacker may tamper with the switches or base stations in order to defeat the location tracking capability. The security of current networks should be improved significantly in order to support network-centric schemes, as witnessed by successful break-ins of phone freaks. We do not address this kind of attack here, and assume that the equipment under the network provider's control is immune from tampering.

Using Cellular Telephony

The cellular telephony infrastructure has upcoming capabilities for locating individual subscribers. Service providers can use these capabilities to locate STTs. In the cellular E911 scheme, the STT contains a *cellular wireless modem*, which is used for return path communications (callbacks) to the service provider. Each time the STT performs a callback, the service provider can use the *wireless enhanced 911* infrastructure to determine the location of the STT. This location information would then be compared to the location where the STT is supposed to be, and if the locations do not match, an appropriate action could be taken.

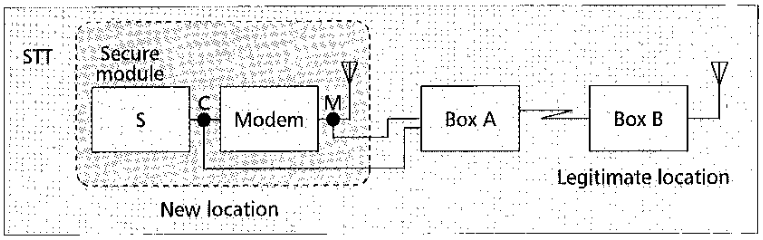
The cost of the cellular modem is less than $100, since it is a bare-bones device, much simpler than a cellular phone. It does not require components such as a display, keypad, microphone, speaker, or vocal encoder. The antenna of the cellular modem may be incorporated into the STT, which would provide adequate communication quality in most residential settings. Otherwise, it may need an external antenna and an amplifier. The added cost to the cellular infrastructure is negligible once wireless E911 is supported by the infrastructure.

Since the STTs share the same infrastructure as real emergency calls, we must ensure that cellular E911 does not interfere with emergency call processing. We also want to reduce the number of additional stationary users the cellular base stations have to support. This can be achieved in several ways. First, the STTs should activate their cellular modems only for callbacks, and the callback times should be staggered. Second, the provider may initiate callbacks (see the next section) by sending coded messages in the satellite broadcast to activate only a subset of the cellular modems. In any case, cellular operators and content providers must reach an agreement to restrict the impact of cellular E911 on the infrastructure.

How Cellular-E911 Works

E911 service has proliferated in the U.S. wireline marketplace. This service provides accurate 911 call routing to the appropriate emergency dispatch center, and relays the caller's location, name, and telephone number to the dispatch operator. In 1996 the FCC developed requirements to provide a similar level of service to wireless users [11].

The FCC ruling proposes a two-phase implementation. In phase 1, scheduled for 18 months beyond the effective date of October 1, 1996, wireless carriers must support call routing based on the caller's cell/sector. In phase 2, scheduled for five



■ Figure 4. *A general antenna-extending attack against the cellular E911 scheme.*

years beyond the effective date, the carriers must support technology to determine a caller's location within 125 m for 67 percent of all wireless 911 calls.

Phase 1 is completely network-centric, and relatively easy to achieve, since the knowledge of the caller's cell/sector exists in the network. However, the accuracy varies with the sector size, which could be as large as a few kilometers in rural areas.

The implementation of phase 2 may be either network-centric or terminal-centric (for a survey see [12]). Terminal-centric solutions require a GPS receiver to be integrated into the mobile phone, which would then report its position. For our purposes, this technology would suffer from the shortcomings of the GPS-based solutions we discuss later. These shortcomings are inherent in all terminal-centric approaches.

Phase 2 can be implemented by several network-centric techniques. All of these techniques determine the caller's location based on measurements from several base stations. The techniques involve measuring the signal strength, the direction or time difference of signal arrival, or the round-trip time between the mobile phone and the base station. For our purposes, though, all of these network-centric techniques are equally suitable and offer a marked improvement in accuracy over the basic phase 1, without the problems of GPS-based solutions. Several vendors offer equipment meeting the FCC accuracy requirements. Field trials are being conducted in various areas, and the initial results are encouraging, although somewhat short of the requirements [13].

Cellular E911 Accuracy and Additional Features

The accuracy of cellular E911 is determined by the accuracy of the cellular location tracking. In the immediate future this would mean an uncertainty region as large as the cell/sector in which the STT is located, an area which can have a diameter measured in kilometers in rural areas. However, this accuracy is expected to improve to about 125 m over the next few years, as described previously.

Cellular E911 has the additional advantage that it does not use the customer's phone line for callbacks. Therefore, the callbacks can be scheduled to achieve optimal load balancing at the service provider's equipment without regard to the customers' phone usage habits. Moreover, the STT does not need to be placed near a phone jack, which simplifies installation and reduces its cost. The importance of this kind of cost reduction in the cable TV industry is attested to by the fact that TCI reportedly plans to use indoor wireless modems (which transmit over the existing 110 V wiring) in their STTs to avoid the costs of installing new phone jacks in the customer's home [14]. Moreover, using a dedicated communication channel (not the telephone wireline) enables the service provider to send short-lived decryption keys to STTs much more frequently (not just once a month), so the damage of breaking one key is smaller.

Another advantage of cellular E911 is the ability to perform provider-initiated callbacks. For instance, the service provider may choose to initiate a callback immediately prior to a popular pay-per-view event. Cellular E911 could then be used to detect the STTs that were illicitly moved, say, from a residence to a bar.

Finally, cellular E911 is the obvious solution for the WLL fraud problem. The WLL unit is essentially a cellular phone, so if the phone company has location tracking capabilities it can easily address this issue.

Cellular E911 Vulnerability

Cellular E911 has the advantage of being network-centric, and hence does not suffer from most of the spoofing attacks to which the GPS scheme is vulnerable. Moreover, a cellular phone cannot masquerade as being in a location other than where its antenna is actually transmitting from. Therefore, as long as the cellular phone infrastructure is not tampered with (e.g., the pirate does not attack the base station equipment), the information conveyed to the service provider identifies where the transmission is coming from.

A simple-minded attack on cellular E911 disconnects the cellular antenna from the STT any time it is activated outside its legitimate location. This would thwart provider-initiated callbacks. However, the STT must communicate with the provider at least once per billing period to receive new keys. During the callback, the STT may report the periods in which it was operating, and the provider may check it against the list of failed provider-initiated callbacks. Another approach is to freeze the STT after a small number of failed callbacks, which the STT initiates at random times when it is operational.

Since the scheme relies on identifying the position of the transmitting antenna, a pirate essentially needs to "extend" the connection between the STT and its cellular antenna. As long as the cellular antenna is in its original location, the movement of the STT will not be detected. In order to move the STT from its legitimate location to a new location, the pirate could build the setup depicted in Fig. 4.

There are several variants of the attack, according to the communication layer the pirate chooses to handle. In one variant, the pirate disconnects the cellular modem from its antenna (at point M in Fig. 4) and attaches box A instead. Box A in turn is connected to box B at the legitimate location using some alternative communication network (e.g., the Internet), and B is connected to a cellular antenna. A and B act as repeaters and transparently relay the signals between the wireless modem at the new location and the antenna at the legitimate one. A naive implementation of this attack would require a relatively high bandwidth between A and B in order to relay the 800 MHz or 1800/1900 MHz analog signals used in cellular communications. A more sophisticated design would have the capability to extract (and recreate) higher levels of the protocol between the wireless modem and the base station, thus reducing the bandwidth requirement between A and B.

In another variant, the pirate disconnects the secure module from the modem (at point C in Fig. 4) and attaches box A there. Box B would then contain a complete cellular modem, and the two boxes would relay the digital information corresponding to the protocol between the the secure module and the modem.

The antenna-extending attack could be detected by imposing strict timing constraints on the communication protocol between the STT and the nearest base station. Any low-cost communication network used by boxes A and B (e.g., the Internet) would introduce delays that could be detected this way.

Using the Global Positioning System

This section describes the GPS scheme, which could be employed anywhere on the globe.

How GPS Works

The GPS uses a constellation of 24 satellites with synchronized clocks that broadcast a specific bitstream. GPS receivers compute their position from the phase shift between the signals received from four or more satellites. In addition to accurate positioning, GPS also provides a fairly accurate clock (340 nanosecond accuracy). An accessible introduction to GPS is [15].

The GPS scheme requires STTs to contain a GPS receiver in addition to the secure module. The GPS receiver requires an additional external antenna, which should be packaged with the satellite receiving dish. If the down-lead from the antenna to the STT is short, there is no need for amplification, and the added cost is negligible. Standard installation should include both antennas, so customers need not be aware of the added complexity. Callbacks would still be performed over the phone network. Prior to the callback, the STT would query the GPS receiver to get its current position. This location information would then be sent to the service provider, who would match it against the STT's legitimate location recorded in the database.

Figure 5 depicts two possible configurations of the STT: In configuration a, the STT consists of a tamper-resistant secure module (S), and an off-the-shelf GPS receiver (G). S deals with unscrambling the broadcast information and communicating with the service provider, and queries G for position information using G's data interface. In configuration b, the STT consists of a single tamper-resistant module (S + G), which combines the functions of S and G of configuration a. Configuration b would be more expensive to design and build, since it would require incorporating part or all of the GPS functionality into the secure module.

GPS Accuracy

Cheap commercial devices can determine their position with 100 m accuracy. Differential GPS (DGPS) can determine the location with 10 m accuracy, but requires an additional known signal source.

GPS Vulnerability

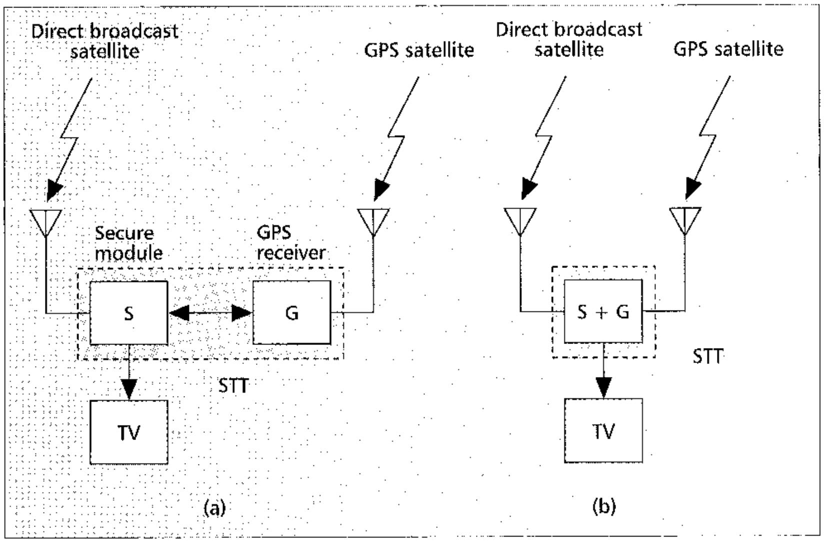
The main problem with the GPS scheme is that it is completely terminal-centric. The service provider relies on the STT to measure its position and report this information correctly. Therefore, a pirate has several options for attacking the STT which would cause it to always report its expected legitimate position regardless of its true whereabouts.

The cheaper configuration, a, can be defeated by various forms of man-in-the-middle attack. If the communication between S and G is in the clear, the pirate can substitute the expected position for the actual one. An obvious countermeasure is to have S and G use a secure protocol between them (with the side effect that G is no longer completely off the shelf). This protocol would require at least the following properties:

- The protocol must mutually authenticate S and G; otherwise, the pirate can disconnect G and replace it by a module that always reports a legitimate location.
- The GPS reports must contain freshness information such as timestamps or some random bits; otherwise, the pirate can record the encrypted report at the legitimate location and replay it at the new location.
- The GPS reports must be signed or have a message authentication code (MAC) attached [16] to prevent pirate tampering.

Even if a secure protocol is used in configuration a, or the STT is designed using configuration b, a sophisticated pirate can still defeat the scheme without tampering with its internals at all. The pirate can generate fake signals from the GPS satellites and feed them into the GPS antenna. Since the trajectory of the GPS satellites is known with great accuracy, the pirate can compute the signal that would be received on any point on the Earth at any given time. Of course, we are assuming here that the GPS receiver uses the Standard Positioning Service (SPS), which is the civilian (nonmilitary) broadcast. If the GPS receiver uses the Precise Positioning Service (PPS), which is currently available only to the U.S. and allied military, the pirate would not be able to generate the signals since they are encrypted.

Another weakness of GPS is that GPS signals may be jammed by a low-power transmitter, as described in recent reports [17–19]. The reports indicate that a single transmitter may inhibit the operation of GPS receivers in a large area or cause major locating errors.



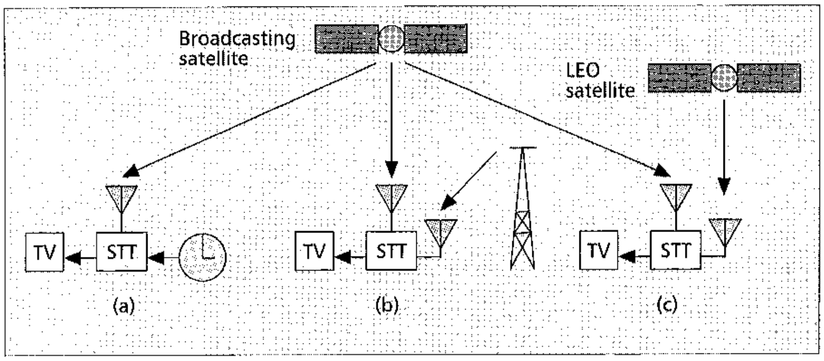
■ Figure 5. *Configurations of an STT design using GPS: a) separate S and G modules; b) combined S and G modules.*

Using Satellite Ranging

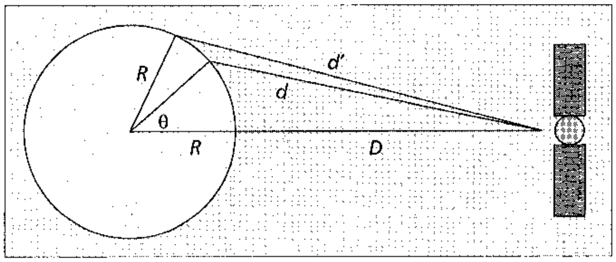
The last scheme we describe is the *Sat-Range* scheme, which determines the STT's location using a technique similar to that of GPS. Sat-Range computes the *distance* between an Earth-bound receiver (the STT in our case) and a satellite, based on the signal's time of arrival. We capitalize on the fact that the STT is already receiving a satellite signal, namely the signal carrying the TV broadcast. Sat-Range measures the phase shift in the satellite signal relative to an external signal, which may be a synchronized clock or a message received from another source, such as a terrestrial paging network or low earth orbit (LEO) messaging system. Figure 6 depicts the operation of Sat-Range.

How Sat-Range Works

The Sat-Range scheme requires the STT to record the next t bits from the satellite broadcast starting at some absolute time



■ Figure 6. *The satellite ranging scheme: a) using a synchronized clock; b) using the paging network; c) using LEO messaging.*



■ Figure 7. *The distance between a point on Earth and a satellite.*

or when it receives an order to start recording. The STT would call the service provider after the recording and report the t bits it has recorded. If the STT has moved away from its legitimate location, the t bits it records would be shifted relative to the bits it would have recorded at its legitimate location. This shift is caused by the difference in the distance between the broadcasting satellite and the STT.

We propose three alternatives for supplying the external signal to start recording. In all three cases the command to start the recording is encrypted, so only the target STT would recognize it. In the first case, depicted in Fig. 6a, we use an accurate clock which shows the same time in all STTs. The service provider broadcasts an encrypted message via the satellite that addresses a subset of the STTs. The message specifies some absolute time in the future for starting the recording of t bits. The bitstreams that different STTs record depend on their distances from the satellite, since all of the STTs start recording at the same instant. We can keep an accurate clock by using a GPS system, or including timing signals in the satellite TV broadcast and sending a correction individually to each STT based on its legitimate location. The correction will be included in the first communication between the STT and the service provider. The STT can derive an accurate clock pulse from the satellite's signal, which is inherently more accurate than an internal low-cost clock.

In the second alternative, depicted in Fig. 6b, we use a terrestrial paging system to send an encrypted request to start recording at the instant the message is received. Another approach is to use a separate geosynchronous satellite paging system, and receive the encrypted message directly from the satellite.

In the third alternative, depicted in Fig. 6c, we use a LEO satellite messaging system, such as the ORBCOMM system [20, 21], to send an encrypted request to start recording at the instant the message is received.

The service provider can detect relative movements of STTs by asking a group of STTs in the same geographical proximity to start recording at the same time. The geographical proximity is derived from the legitimate location of the corresponding STTs. If the reported recording from one of the STTs is shifted significantly relative to the others, it would indicate that this STT has been moved. Another possibility is to include a trusted STT in each of those groups. The trusted STT would be located on a utility pole, say, so its position is always trusted.

The cost of the STT is not increased at all if it uses an accurate clock based on broadcast time signals and a per-STT correction. The cost of a scaled-down GPS receiver used for obtaining an accurate time is less than $100. The cost of a paging receiver is less than the cost of a full-featured pager; we estimate it to be less than $25. The cost of a LEO message receiver should be low, probably less than $100, since this technology is designed to be used in low-cost high-volume applications such as remote meter reading. A side benefit of using LEO for communication is that the communication can be bidirectional, so there is no need to use the customer's phone line.

Sat-Range Accuracy

Consider a digitally modulated signal arriving from a particular transponder on the satellite at baud rate r. The raw bit rate would then be rk b/s, where the ratio k depends on the modulation and error correction schemes used in the physical layer. For instance, the DirecTV system [22] uses quadrature phase shift keying (QPSK) modulation [23,24] at a baud rate of 27 MHz to achieve a bit rate of ≈ 30 Mb/s (i.e., $k \approx 1.11$).

The movement of an STT can be detected only if it results in a phase shift of more than one cycle. Assume that the STT's distance from the satellite changes from d to d' and the signal is propagated at the speed of light c. Then a phase shift in the digital bitstream can be detected only if

$$\frac{|d'-d|}{c} \geq \frac{1}{r}.$$

Let $s = c/r$ denote the minimal distance change that will cause a detectable phase shift. Then a movement from distance d to distance d' would result in a shift of $k\lfloor |d'-d|/s \rfloor$ bits in a k-bit-per-baud modulation scheme. For instance, in the DirecTV example, a change of about 11 m in distance to the satellite would cause a 1-bit shift.

The distance d between an STT on Earth and a geosynchronous satellite is determined by several parameters. Assume that the STT is located at longitude λ_1, latitude θ, and altitude h, and the satellite is at longitude λ_2 above the equator at the geosynchronous altitude of $D = 35{,}803$ km. Let $\lambda = |\lambda_1 - \lambda_2|$ be the difference between the longitudes. Then the distance to the satellite (also called *slant range*) is

$$d = \sqrt{(R+h)^2 + (R+D)^2 - 2(R+h)(R+d)\cos\theta\cos\lambda},$$

where $R = 6367$ km is the radius of the Earth. Figure 7 depicts the measurement of the slant range. For simplicity, the figure assumes that the STT is at altitude $h = 0$, and the satellite and STT are on the same longitude (i.e., $\lambda = 0$).

Using the above formula, we can compute the magnitude of *terrestrial* movement that would cause a change of s in the distance to the satellite. For instance, in the DirecTV example, the satellite is located at longitude λ_2 = 101°W. Using this value, our computations show that a movement of 15–20 m either north or south anywhere in the continental United States would then be noticeable.

However, this computation is correct only if the STT has a perfect clock, which is unrealistic. If the STT's clock has an accuracy of a ns, an observed change of ac m in the distance to the satellite can be a measurement error (where c is the speed of light). For instance, if the clock is accurate up to 340 ns, as is the case for the SPS signal from GPS, the system would only detect differences of at least 102 m in the distance to the satellite. In the DirecTV example, our computations show that terrestrial north-south movements of 140–180 m in the United States would be detectable. Specifically, moving 146 m north from Buffalo, New York (43°N, 79°W) or 176 m south from San Diego, California (32°N, 117°W) should be detectable. We selected Buffalo and San Diego since these cities are close to the Canadian and Mexican borders, respectively.

Note that since we are ranging to a single satellite, the points which are at distance d from the satellite define the surface of a sphere centered at the satellite. As long as the STT stays on this sphere, no phase shift will occur. For instance, if the new location has the same altitude as the old, the points that are equidistant from the satellite typically define a circle on Earth, which means that there are two directions of movement that can go undetected (along this circle). However, given that the technique's sensitivity is in the 140–180 m range, we argue that moving the STT without changing its distance to the satellite is a rather unlikely event.

Sat-Range Strengths and Vulnerabilities

The Sat-Range scheme is based on encrypted messages that are sent from the service provider to the STTs. An eavesdropper cannot recognize when the service provider asks a particular STT to start recording. Only when the STT contacts the service provider (some time after the recording has ended) can the eavesdropper discover that a recording has been made.

If Sat-Range uses a synchronized clock, the STT can be moved to another location with the same distance to the satellite without being detected. As noted above, since Sat-Range is sensitive enough to detect a 180 m terrestrial move, it is extremely unlikely to accomplish this.

However, in some cases the synchronized clock can be defeated by adding a buffer between the receiving antenna and the STT. The buffer compensates for moving the STT *closer* to the satellite by delaying the incoming signal by a fixed amount. The delay amount is equivalent to the difference in signal propagation time between the new location and the legitimate location. However, the buffer cannot help if the STT is moved *away* from the satellite.

If Sat-Range uses a terrestrial paging system or geosynchronous satellite paging to send the recording request, the pirate can use buffering to compensate for the STT movement. Since the STT starts recording immediately when it receives a paging message, the pirate intercepts and delays every paging message by a fixed amount if the paging message arrives ahead of time due to the movement. If the paging message arrives later due to the movement, the pirate delays the satellite broadcast by a fixed amount. Note that this delay can be computed in advance, since the locations of the broadcast satellite and the paging request source (either a terrestrial system or geosynchronous satellite paging) are known in advance. However, buffers for high-bit-rate signals are likely to be expensive and require modification of the STT.

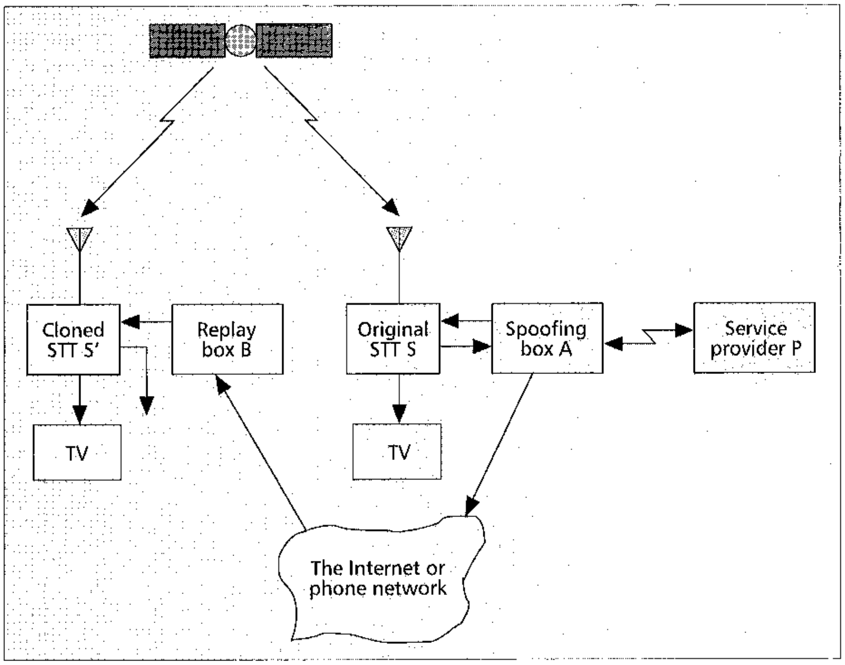
If Sat-Range uses a LEO messaging system, the distance between the STT and the LEO satellites constantly changes. Moreover, since the instant when the LEO satellite sends the recording request cannot be predicted, the signal propagation delay between the LEO and the STT cannot be precomputed and varies from one request to the next. Thus, the buffering approach will fail to hide a movement of the STT. Another strength of this technique is that there are no "safe" zones, as in the synchronized clock case, that can hide movement.

Complete Cloning

All of the previously-mentioned schemes for detecting the STT's location will fail if one of the STTs is completely cloned. By this we mean that the cloned STT is identical to the original STT, including all the secret keys and individual identification of the original box. In addition, the internal state of the cloned STT can be kept identical to that of the original STT if it receives the same replies at the same time from the service provider as the original STT. Figure 8 depicts a system that can support an unlimited number of cloned STTs. Note that cloning can be achieved by brute-force reverse engineering without actual knowledge of the protocols and algorithms inside the cloned STT.

The system works as follows. The original STT S is connected to spoofing box A, which collects and relays the messages service provider P sends to S. Replay box B receives these messages and feeds them to the cloned STT S′. The messages S′ sends are discarded. A and B exchange synchronization information which ensures that S and S' are synchronized. The communication between A and B may be done via the Internet or through a phone call.

Since A does not alter the communication between S and P, and does not introduce long delays, P cannot detect the presence of A, so it will send new keys to S. Since S and S′ implement the same communication algorithm, share the internal state, and operate in sync, the messages P sends to S will also be accepted by S′, including the



■ Figure 8. *Complete cloning of an STT.*

messages that contain new keys. Moreover, even if S′ initiates the call to P, S will do the same at the same time, since they have the same internal state. S and S′ will behave differently only if they contain a hardware random number generator that employs a physical phenomenon to generate random numbers.

Since all the messages sent by S′ are discarded, P is not aware of the existence of S′, so P cannot detect whether S′ is located in a legitimate location or not.

Complete synchronization between A and B is needed only when the protocol between S and P is dependent on the current time. If it is not, an offline replay of the messages sent from P to S is sufficient to update the state of S′.

Summary and Conclusions

We presented four schemes — caller ID, cellular E911, GPS, and Sat-Range — that allow service providers to monitor the location of customers' equipment using existing infrastructure. These schemes can be used to detect the unauthorized movement of supposedly stationary equipment, such as set-top boxes, wireless local loop phones, and surveillance equipment. Currently, the most common use of location tracking is in the direct broadcast satellite industry, where many service providers use caller ID to locate their customers. The other three schemes are practical and inexpensive to build and use, and are more robust than the caller ID scheme. Nevertheless, a sufficiently motivated pirate would probably be able to circumvent them. Therefore, we do not consider the problem to be completely solved. It is still a very interesting research question to come up with other schemes that may offer even better protection.

Acknowledgments

We are grateful to Dan Heer for introducing us to the problems of direct broadcast satellites, and for many valuable discussions. Dan Heer was first to suggest the idea of using the E911 infrastructure. We thank Bob Richton for many useful discussions regarding wireless local loop technology. A preliminary version of this article appeared in [25].

References

[1] R. Anderson and M. Kuhn, "Low Cost Attacks on Tamper Resistant Devices," *5th Security Protocols Wksp., Lecture Notes in Comp. Sci.*, vol. 1361, Paris, France, Apr. 1997, pp. 125–36.

[2] J. McCormac, *European Scrambling Systems 5*, Waterford Univ. Press, 1996.

[3] B. M. Macq and J.-J. Quisquater, "Cryptology for Digital TV Broadcasting," *Proc. IEEE*, vol. 83, no. 6, 1995, pp. 944–57.

[4] A. Wool, "Key Management for Encrypted Broadcast," *Proc. 5th ACM Conf. Comp. and Commun. Security*, San Francisco, CA, Nov. 1998, pp. 7–16.

[5] J.-J. Quisquater, Personal communication, 1998.

[6] M. Kibati and D. Krairit, "The Wireless Local Loop in Developing Regions," *Commun. ACM*, vol. 42, no. 6, June 1999, pp. 60–66.

[7] A. R. Noerpel and Y.-B. Lin, "Wireless Local Loop: Architecture, Technologies, and Services," *IEEE Pers. Commun.*, June 1998, pp. 74–80.

[8] Automatic Number Identification (ANI) and Operator Number Identification (ONI), FSD 20-20-0000, LATA Switching Systems Generic Requirements (LSSGR), Bellcore Tech. Ref. TR-NWT-000682, issue 2, Mar. 1991.

[9] CLASS Feature: Calling Number Delivery, FSD 01-02-1051, LATA Switching Systems Generic Requirements (LSSGR), Bellcore Tech. Ref. TR-NWT-000031, issue 4, Dec. 1992.

[10] M. Kuhn, *Personal Communication*, 1999.

[11] "Report and Order and Further Notice of Proposed Rulemaking in the Matter of Revision of the Commission's Rules to Ensure Compatibility with Enhanced 911 Emergency Calling Systems," FCC Docket No. 94-102, June 1996.

[12] M. J. Meyer *et al.*, "Wireless Enhanced 911 Service–Making It a Reality," *Bell Labs Tech. J.*, Autumn 1996, pp. 188–202.

[13] "Report on the New Jersey Wireless Enhanced 911 System Trial January 22 to April 30, 1997: The First 100 Days," State of NJ Dept. of Law and Public Safety; Div. of State Police, June 1997; available at http://www.trueposition.com/pressclip6.htm.

[14] P. J. Britt, "Return Path Simplified: Wireless Modem Jacks Save Cable TV installation time," *Telephony*, vol. 234, no. 7, Feb. 1998, p. 14.

[15] P. H. Dana, "Global Positioning System Overview," The Geographer's Craft Project, Dept. of Geography, Univ. TX Austin, 1998, http://www.utexas.edu/depts/grg/gcraft/notes/gps/gps.html

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

[17] B. Brewin, "Rogue Transmitter Knocks Out GPS Signals," *Fed. Comp. Week*, Apr. 13, 1998; http://www.fcw.com/pubs/fcw/1998/0413/fcw-frontgps-4-13-1998.html

[18] O. Schmidt, "GPS-Jammers Too Good for Their Own Good," *Intelligence*, vol. 76, no. 6, 1998.

[19] aims@ext.jussieu.fr (AIMS/Intel-Info), "Two Known GPS Jamming Cases," *RISKS Forum Digest*, vol. 19, no. 74, May 1998; available at http://catless.ncl.ac.uk/Risks/19.74.html and ftp://unix.sri.com/risks/risks-19.74

20] I. Brodsky, "Will They Fly or Crash 'n' Burn?, *Telephony*, Apr. 27 1998, pp. 34–42.

[21] ORBCOMM Corp; http://www.orbcomm.com/about/sysdesc.html

[22] "The High-Tech Behind Broadcasting DirecTV," DirecTV Inc., 1998; http://www.directv.com/hardware/tech.html

[23] T. T. Ha, *Digital Satellite Communications*, 2nd ed., McGraw-Hill, 1990.

[24] ITU-R BO.1211, "Digital Multi-Programme Emission Systems for Television, Sound and Data Services for Satellites Operating in the 11/12 GHz Frequency Range," 1995.

[25] E. Gabber and A. Wool, "How to Prove Where You Are: Tracking the Location of Customer Equipment," *Proc. 5th ACM Conf. Comp. and Commun. Security*, San Francisco, CA, Nov. 1998, pp. 142–49.

Biographies

ERAN GABBER (eran@research.bell-labs.com) received his Ph.D. in computer science from Tel-Aviv University, Israel, in 1993. From 1993 to 1995 he was the principal researcher at nSOF Parallel Software, Ltd., a startup company in Israel. He joined Bell Laboratories in 1995, where he is a member of technical staff in the Information Sciences Research Center. His research interests include operating systems, quality of service, Internet security, and electronic commerce.

AVISHAI WOOL (yash@research.bell-labs.com) received a B.Sc. in mathematics and computer science from Tel-Aviv University, Israel, in 1989, and an M.Sc. and Ph.D. in computer science from the Weizmann Institute of Science, Israel, in 1992 and 1996, respectively. He then joined Bell Laboratories, where he is a member of technical staff. His research interests include distributed computing, quorum systems, security, and fast communication networks.