Location-Aware and Safer Cards: Enhancing RFID Security and Privacy via Location Sensing

Di Ma, Member, IEEE, Nitesh Saxena, Member, IEEE, Tuo Xiang, and Yan Zhu

Abstract—In this paper, we report on a new approach for enhancing security and privacy in certain RFID applications whereby location or location-related information (such as speed) can serve as a legitimate access context. Examples of these applications include access cards, toll cards, credit cards, and other payment tokens. We show that location awareness can be used by both tags and back-end servers for defending against unauthorized reading and relay attacks on RFID systems. On the tag side, we design a location-aware selective unlocking mechanism using which tags can selectively respond to reader interrogations rather than doing so promiscuously. On the server side, we design a location-aware secure transaction verification scheme that allows a bank server to decide whether to approve or deny a payment transaction and detect a specific type of relay attack involving malicious readers. The premise of our work is a current technological advancement that can enable RFID tags with low-cost location (GPS) sensing capabilities. Unlike prior research on this subject, our defenses do not rely on auxiliary devices or require any explicit user involvement.

Index Terms—RFID, mobile payment system, relay attacks, context recognition, location sensing

1 INTRODUCTION

Low cost, small size, and the ability of allowing computerized identification of objects make Radio Frequency IDentification (RFID) systems increasingly ubiquitous in both public and private domains. Prominent RFID applications supply chain management (inventory control) [16], e-passports [57], credit cards [15], driver’s licenses [60], [41], vehicle systems (toll collection or car key) [17], [27], [25], access cards (building, parking or public transport) [46], and medical implants [38]. NFC, or Near Field Communication [26], is yet another upcoming RFID technology that allows devices, such as smartphones, to have both RFID tag and reader functionality. In particular, the use of NFC-equipped mobile devices as payment tokens (such as Google Wallet) is considered to be the next generation payment system and the latest buzz in the financial industry [10].

A typical RFID system consists of tags, readers, and/or back-end servers. Tags are miniaturized wireless radio devices that store information about their corresponding subject. Such information is usually sensitive and personally identifiable. For example, a US e-passport stores the name, nationality, date of birth, digital photograph, and (optionally) fingerprint of its owner [29]. Readers broadcast queries to tags in their radio transmission ranges for information contained in tags and tags reply with such information. The queried information is then sent to the server (which may coexist with the reader) for further processing and the processing result is used to perform proper actions (such as updating inventory, opening gate, charging toll or approving payment).

Due to the inherent weaknesses of underlying wireless radio communication, RFID systems are plagued with a wide variety of security and privacy threats [28]. A large number of these threats are due to the tag’s promiscuous response to any reader requests. This renders sensitive tag information easily subject to unauthorized reading [23]. Information (might simply be a plain identifier) gleaned from a RFID tag can be used to track the owner of the tag, or be utilized to clone the tag so that an adversary can impersonate the tag’s owner [28].

Promiscuous responses also incite different types of relay attacks. One class of these attacks is referred to as “ghost-and-leech” [34]. In this attack, an adversary, called a “leech,” relays the information surreptitiously read from a legitimate RFID tag to a colluding entity known as a “ghost.” The ghost can then relay the received information to a corresponding legitimate reader and vice versa in the other direction. This way a ghost and leech pair can succeed in impersonating a legitimate RFID tag without actually possessing the device.

A more severe form of relay attacks, usually against payment cards, is called “reader-and-ghost”; it involves a malicious reader and an unsuspecting owner intending to make a transaction [14]. 1 In this attack, the malicious reader, serving the role of a leech and colluding with the ghost, can fool the owner of the card into approving a transaction which she did not intend to make (e.g., paying for a diamond purchase made by the adversary while the owner only intending to pay for food). We note that addressing this problem requires secure transaction verification, i.e., validation that the tag is indeed authorizing the intended payment amount.

1. In contrast to the “ghost-and-leech” attack, the owner in the “reader-and-ghost” attack is aware of the interrogation from the (malicious) reader.
The feasibility of executing relay attacks has been demonstrated on many RFID (or related) deployments, including the Chip-and-PIN credit card system [14], RFID-assisted voting system [42], and keyless entry and start car key system [17].

With the increasingly ubiquitous deployment of RFID applications, there is a pressing need for the development of security primitives and protocols to defeat unauthorized reading and relay attacks. However, providing security and privacy services for RFID systems presents a unique and formidable set of challenges. The inherent difficulty stems partially from the constraints of RFID tags in terms of computation, memory and power, and partially from the unusual usability requirements imposed by RFID applications (originally geared for automation). Consequently, solutions designed for RFID systems need to satisfy the requirements of the underlying RFID applications in terms of not only efficiency and security, but also usability.

1.1 Sensing-Enabled Automated Defenses

Although a variety of security solutions exist, many of them do not meet the constraints and requirements of the underlying RFID applications in terms of (one or more of) efficiency, security, and usability. We review related prior work in Section 2.

In an attempt to address these drawbacks, this paper proposes a general research direction—one that utilizes sensing technologies—to address unauthorized reading and relay attacks in RFID systems without necessitating any changes to the traditional RFID usage model, i.e., without incorporating any explicit user involvement beyond what is practiced today. The premise of the proposed work is based on a current technological advancement that enables many RFID tags with low-cost sensing capabilities. Various types of sensors have been incorporated with many RFID tags [48], [24], [49], Intel’s Wireless Identification and Sensing Platform (WISP) [50], [54] is a representative example of a sensor-enabled tag, which extends RFID beyond simple identification to in-depth sensing. This new generation of RFID devices can facilitate numerous promising applications for ubiquitous sensing and computation. They also suggest new ways of providing security and privacy services by leveraging the unique properties of the physical environment or physical status of the tag (or its owner). In this paper, we specifically focus on the design of context-aware security primitives and protocols by utilizing sensing technologies so as to provide improved protection against unauthorized reading and relay attacks.

The physical environment offers a rich set of attributes that are unique in space, time, and to individual objects. These attributes—such as temperature, sound, light, location, speed, acceleration, or magnetic field—reflect either the current condition of a tag’s surrounding environment or the condition of the tag (or its owner) itself. A sensor-enabled RFID tag can acquire useful contextual information about its environment (or its owner, or the tag itself), and this information can be utilized for improved RFID security and privacy without undermining usability.

1.2 Our Contributions

In this paper, we report on our work on utilizing location information to defend against unauthorized reading and relay attacks in certain applications. We notice that in quite some applications, under normal circumstances, tags only need to communicate with readers at some specific locations or while undergoing a certain speed. For example, an access card to an office building needs to only respond to reader queries when it is near the entrance of the building; a credit card should only work in authorized retail stores; toll cards usually only communicate with toll readers in certain fixed locations (toll booths) or when the car travels at a certain speed. Hence, location or location-specific information can serve as a good means to establish a legitimate usage context.

Specifically, we present two location-aware defense mechanisms for enhanced RFID security and privacy. First, we show that location information can be used to design selective unlocking mechanisms so that tags can selectively respond to reader interrogations. That is, rather than responding promiscuously to queries from any readers, a tag can utilize location information and will only communicate when it makes sense to do so, thus, raising the bar even for sophisticated adversaries without affecting the RFID usage model. For example, an office building access card can remain locked unless it is aware that it is near the (fixed) entrance of the building. Similarly, a toll card can remain locked unless the car is at the toll booth and/or it is traveling at a speed range regulated by law.

Second, we show that location information can be used as a basis for secure transaction verification to defend against the reader-and-ghost attacks, a devastating relay attack against mobile payment systems involving malicious readers. This is based on a straightforward observation that, under normal scenarios, both the legitimate tag and legitimate reader are in close physical proximity, at roughly the same location. Thus, if the two devices indicate different physically disparate locations, a bank server could detect the presence of a reader-and-ghost attack. For example, the bank server can deny the transaction when it detects the valid tag (RFID credit card) is located in a restaurant, while the valid reader is attack presented in a jewelry shop and prevent the attack presented in [14].

For deriving location information, we make use of the well-known global positioning system (GPS). To demonstrate the feasibility of our location-aware defense mechanisms, we first integrate a low-cost GPS receiver with a RFID tag (the Intel’s WISP), and then conduct relevant experiments to acquire location and speed information from GPS readings. Our experimental results show that it is possible to measure location and speed with high accuracies even on a constrained GPS-enabled platform, and that our location-aware defenses are quite effective in thwarting many attacks on RFID systems without affecting the RFID usage model. Besides the traditional RFID tags, our location-aware defenses are also directly applicable to NFC-enabled phones, which often come readily equipped with GPS receivers.

Scope of our work. Privacy of users is very important in RFID applications but at the same an objective and realistic definition of privacy is needed. This means that the level of privacy assured in an RFID system can be no more than what is provided in the real world. In the real world, a vehicle can be identified by its registration plate and then
tracked by road authorities via video cameras as is the practice nowadays. Similarly, a credit card can be tracked by its issuing bank whenever a transaction is made. Hence, the objective of our privacy protection is to prevent privacy leakage due to unauthorized parties.

We also note that, in some applications, the proposed approaches may not provide absolute security. However, they still significantly raise the bar even for sophisticated adversaries without affecting the RFID usage model. For example, the selective unlocking mechanism for toll cards, based solely on speed detection, will leave the card vulnerable in other situations where the car is undergoing the same speed designated at the toll booths. However, it still protects the car from being read by an adversary while traveling at other speeds or when stationary. In addition, although the proposed techniques can work in a stand-alone fashion, they can also be used in conjunction with other security mechanisms, such as cryptographic protocols, to provide stronger cross-layer security protection.

1.3 Economic Feasibility

A fundamental question with respect to our sensing-enabled approaches is whether the cost of sensor-enabled tags is acceptable. The cost of an RFID tag is dependent on several factors such as the capabilities of the tag (computation, memory), the packaging of the tag (e.g., encased in plastic or embedded in a label), and the volume of tags produced. High-end RFID tags, such as those available on e-passports or some access cards that are capable of performing certain cryptographic computations, cost around $5; whereas low-end inventory tags that do not support any (cryptographic) computation cost only about $0.20 [58]. (We emphasize that our proposal generally targets high-end RFID tags that open up a wide array of applications and generally require higher level of security and privacy. Inventory tags, at least for the time being, are not within the scope of our research.)

The current cost of WISP tags—equipped with a thermometer and an accelerometer—assembled from discrete components is roughly $25, but it is expected that this number will be reduced closer to $1 once the WISPs are mass manufactured [9].

Integrating a GPS sensor with an RFID tag is also quite feasible economically. A few GPS-enabled RFID tags have been reported previously. A tag from Numerex and Savi Technology has been equipped with GPS sensors and has the ability to conduct satellite communications [19]. Researchers in Oak Ridge National Laboratory also worked with RFID system suppliers in developing new intelligent tags by combining GPS and environmental sensors [8]; these tags are designed to track goods anywhere within a global supply chain. We note that usually cost of sensing hardware varies greatly not only between different types of sensors but also between various models of the same kind. GPS receivers, in particular, can be as costly as several hundred dollars [55] or as inexpensive as a couple of dollars when purchased in bulk [3]. The estimated cost for the latter is certainly acceptable for high-end tags and does not affect their business model. Incorporating sensors on tags—i.e., increasing the capabilities of tags—may raise the price of tags initially. However, in the long run, following Moore’s law, advances in process technology and mass production should enable tags with more capabilities (such as sensing, increased computation, and memory) at the same cost of today’s tags [12].

Paper outline. The rest of the paper is organized as follows: In Section 2, we review the most relevant prior work on RFID selective unlocking and transaction verification and also provide background information on the current mobile payment infrastructure. Next, we describe our adversary models in Section 3. We present the two proposed location-aware defense mechanisms in Sections 4 and 5, respectively. In Sections 6 and 7, we discuss the design and implementation of our mechanisms, and present our experimental results, respectively. Finally, we discuss related issues that may rise in practice in Section 8, and Section 9 concludes the paper.

2 BACKGROUND AND PRIOR WORK

In this section, we review existing countermeasures against unauthorized reading and relay attacks. We also provide background information about the current mobile payment system which is susceptible to the reader-and-ghost relay attack.

2.1 Prior Work

Hardware-based selective unlocking. Hardware-based selective unlocking schemes have been proposed previously. These include: Blocker Tag [30], RFID Enhancer Proxy [31], RFID Guardian [47], and Vibrate-to-Unlock [39].

All of these approaches, however, require the users to carry an auxiliary device. In Blocker Tag [30], a special RFID tag, called “blocker,” is used to disrupt the identification process used by the reader to identify tags in proximity [30]. RFID Enhancer Proxy [31] and RFID Guardian [47] are special RFID-enabled devices that could be implemented in a PDA or cellphone. They are assumed to come with greater computation capability and, thus, can perform more sophisticated interactions with readers, on behalf of tags, for various security purposes. In Vibrate-to-Unlock [39], a user unlocks his/her RFID tags by authenticating to these tags through a vibrating phone. However, such an auxiliary device (required by above schemes) may not be available at the time of accessing RFID tags, and users may not be willing to always carry these devices.

A Faraday cage can also be used to prevent an RFID tag from responding promiscuously by shielding its transmission. However, a special-purpose cage (a foil envelope or a wallet) would be needed and the tag would need to be removed from the cage in order to be read. This greatly decreases the usability of such solutions as users may not be willing to put up with any changes to the traditional usage model. Moreover, building a true Faraday Cage that shields all communication is known to be a significant challenge. For example, a crumpled sleeve is shown to be ineffective for shielding purposes [36].

In contrast to above schemes, our work does not require the user to carry an auxiliary device or necessitates any additional user involvement. The proposed sensing-enabled automated defense mechanism utilizes sensors present on emerging RFID devices with low-cost sensing capabilities. These sensors, originally deployed not for
security purposes, might be readily available on the RFID devices (such as an NFC-enabled phone) or they could be integrated to an RFID device specially for security purpose with a little cost.

**Cryptographic protocols.** Cryptographic reader-to-tag authentication protocols could also be used to defend against unauthorized reading. However, due to their computational complexity and high-bandwidth requirements, many of these protocols are still unworkable even on high-end tags [28]. There has been a growing interest in the research community to design lightweight cryptographic mechanisms (e.g., [32], [7], [33], [18]). However, these protocols usually require shared key(s) between tags and readers, which is not an option in some applications.

In contrast, our sensing-based defense mechanism is not cryptography-based and, thus, does not require secure association between tags and readers. However, as mentioned Section 1, the proposed techniques can be used in conjunction with cryptographic protocols to provide stronger cross-layer security protection.

**Distance bounding protocols.** These protocols have been used to thwart relay attacks [14], [17]. A distance bounding protocol is a cryptographic challenge-response authentication protocol. Hence, it requires shared key(s) between tags and readers as other cryptographic protocols. Besides authentication, a distance bounding protocol allows the verifier to measure an upper bound of its distance from the prover [6]. (We stress that normal “non-distance-bounding” cryptographic authentication protocols are completely ineffective in defending against relay attacks.) Using this protocol, a valid RFID reader can verify whether the valid tag is within a close proximity thereby detecting ghost-and-leech and reader-and-ghost relay attacks [14], [17]. The upper bound calculated by an RF distance bounding protocol, however, is very sensitive to processing delay (the time used to generate the response) at the prover side. This is because a slight delay (of the orders of a few nanoseconds) may result in a significant error in distance bounding. Because of this strict delay requirement, even XOR- or comparison-based distance bounding protocols [6], [21] are not suitable for RF distance bounding since simply signal conversion and modulation can lead to significant delays. By eliminating the necessity for signal conversion and modulation, a very recent protocol, based on signal reflection and channel selection, achieves a processing time of less than $1\text{ns}$ at the prover side [45]. However, it requires specialized hardware at the prover side due to the need for channel selection. This renders existing protocols currently infeasible for even high-end RFID tags.

**Context-aware selective unlocking.** “Secret Handshakes” is a recently proposed interesting selective unlocking method that is based on context awareness [12]. To unlock an accelerometer-equipped RFID tag [50], [54] using Secret Handshakes, a user must move or shake the tag (or its container) in a particular pattern. For example, the user might be required to move the tag parallel with the surface of the RFID reader’s antenna in a circular manner. A number of unlocking patterns were studied and shown to exhibit low error rates [12]. A central drawback to Secret Handshakes, however, is that a specialized movement pattern is required for the tag to be unlocked. This requires subtle changes to the existing RFID usage model. While a standard, insecure RFID setup only requires users to bring their RFID tags within range of a reader, the Secret Handshakes approach requires that users consciously move the tag in a certain pattern. This clearly undermines the usability of this approach.

“Motion detection” [51] has been proposed as another selective unlocking scheme. Here, a tag would respond only when it is in motion instead of doing so promiscuously. In other words, if the device is still, it remains silent. Although motion detection does not require any changes to the traditional usage model and raises the bar required for a few common attacks to succeed, it is not capable of discerning whether the device is in motion due to a particular gesture or because its owner is in motion. Hence, the false unlocking rate of this approach is high.

In our work, we aim to design location-aware secure RFID schemes that 1) have both low false locking and false unlocking rates, and 2) do not necessitate any changes to the current usage model.

### 2.2 Mobile Payment Infrastructure

EMV, named after its creators, Europay, Mastercard, and Visa, is a global standard for debit and credit card payments. Payment systems based on EMV have been introduced across the world, known by a variety of different names such as “Chip and PIN” [14]. Mastercards PayPass is another EMV compatible “contactless” payment protocol. Fig. 1 presents a simplified version of the EMV-based mobile payment system with online verification. The system consists of three entities of interest: RFID-enabled payment card, the merchant, and the issuer bank, which issues the card. The payment card stores card details such as the credit card number, name of the owner, and expiration date. It also stores a symmetric key shared with its issuer bank. The point-of-sale (PoS) terminal at the merchant side is equipped with an RFID reader. A transaction starts with the merchant issuing a challenge to the payment card. The card calculates a cryptographic response based on the challenge and other information using the key shared with the issuer bank. It then transfers
the response to the merchant terminal through the RFID communication interface. The response is next forwarded by the terminal to the issuer bank, which verifies the response and approves the transaction, if authentication is successful. Our proposed secure transaction verification based on location sensing can work under the current payment infrastructure.

3 Adversarial Models

Our proposed techniques are meant to defend against unauthorized reading, ghost-and-leech, and reader-and-ghost attacks. Adversary models used in the three attack contexts are slightly different. In the following description, we call the tag (reader) under attack as valid tag (reader) and call the tag (reader) controlled by the adversary as malicious tag (reader).

In unauthorized reading, the adversary has direct control over a malicious reader. The malicious reader can be in the communication range of the victim tag without being detected or noticed and, thus, can surreptitiously interrogate the tag. The goal of the adversary is to obtain tag specific information and (later) use such information to compromise user privacy (through inventory checking), clone the tag (and thus impersonate the user), or track the user.

In ghost-and-leech attack, besides the malicious reader (the leech), the adversary has further control over a malicious tag (the ghost), which communicates with a valid reader. The adversary’s goal is to use the malicious tag to impersonate the valid tag by letting the malicious tag respond to interrogations from the valid reader with information surreptitiously read from the valid tag by the malicious reader.

In reader-and-ghost attack, originally called the “mafia fraud” attack [13, 14], the adversary controls a malicious reader and tag pair, just like in the ghost-and-leech attack. However, the malicious reader controlled by the reader-and-ghost adversary is a legitimate reader or believed by the valid tag as a legitimate reader. Hence, the valid tag (or its owner) is aware of and agree with communications with the malicious reader. That is, the interrogation from the malicious reader to the valid tag is not surreptitious as in unauthorized reading and ghost-and-leech attacks. The goal of the adversary is still to impersonate the valid tag.

In all the attack contexts, we assume the adversary does not have direct access to the valid tag, so tampering or corrupting the tag physically is not possible or can be easily detected. The adversary is also unable to tamper the tag remotely through injected malicious code. We further assume that the adversary is able to spoof the GPS signal around the victim tag but not around the victim reader. This is because the reader is usually installed in a controlled place (toll booth, office building gate, or retail store) and, thus, GPS spoofing around the victim reader can be easily detected. We do not consider loss or theft of tags.

4 Location-Aware Selective Unlocking

In this section, we present our location-aware selective unlocking mechanism. It can be used to protect against unauthorized reading and ghost-and-leech attacks. Using location-aware selective unlocking, a tag is unlocked only when it is in an appropriate (prespecified) location. This mechanism is suitable for applications where reader location is fixed and well known in advance. One example application is RFID-based building access system. An access card to an office building needs to only respond to reader queries when it is near the entrance of the building.

A prerequisite in a location-aware selective unlocking scheme is that a tag needs to store a list of legitimate locations \(L_{\text{stored}}\) beforehand (as depicted in Fig. 2). Upon each interrogation from a reader, the tag obtains its current location information \(L_{\text{GPS}}\) from its on-board GPS sensor, and compares it with the list of legitimate locations and decides whether to switch to the unlocked state or not. Due to limited on-board storage (e.g., the WISP has a 8 KB of flash memory) of tags, the list of legitimate locations must be short. Otherwise, testing whether the current location is within the legitimate list may cause unbearable delay and affect the performance of the underlying access system. Moreover, the list of legitimate locations should not change frequently because otherwise users will have to do extra work to securely update the list on their tags. Thus, selective unlocking based on pure location information is more suitable for applications where tags only need to talk with one or a few readers, such as building access cards. It may not be suitable for credit card applications, as there is a long list of legitimate retailer stores, and store closing and new store opening occur on a frequent basis.

Selective unlocking based on pure location information presents similar problems for toll systems as for the credit card systems because toll cards will need to store a long list of toll booth locations. We notice that vehicles mounted with RFID toll tags are usually required to travel at a certain speed when they approach a toll booth. For example, three out of eight toll lanes on the Port Authority’s New Jersey-Staten Island Outer Bridge Crossing permit 25-mph speeds for EZ-Pass drivers; the Tappan Zee Bridge toll plaza, and New Rochelle plaza, NY has 20-mph roll-through speed; Dallas North Toll way has roll-through lanes allowing speeds up to 30 mph. Hence, “speed” can be used as a valid context to design selective unlocking mechanisms for toll cards. That is, a toll card remains in a locked state except when the vehicle is traveling at a designated speed near a toll plaza.

2. In some countries, toll-collection companies have set up roaming arrangements with each other. This permits the same vehicle to use another operator’s toll system, thus reducing setup costs and allowing even broader use of these systems [1].
toll booth (such as 25-35 mph in the Dallas North Toll Way case). GPS sensors can be used to estimate speed either directly from the instantaneous Doppler speed or directly from positional data differences and the corresponding time differences [11].

For better protection against attacks, the speed and location can also be used together as a valid context for unlocking of toll cards. Here, the adversary will only be able to unlock the tag if both the valid location and speed criteria are satisfied.

5 LOCATION-AWARE TRANSACTION VERIFICATION

A highly difficult problem arises in situations when the reader, with which the tag (or its user) engages in a transaction, itself is malicious. For example, in the context of an RFID credit card, a malicious reader can fool the user into approving for a transaction whose cost is much more than what she intended to pay. That is, the reader terminal would still display the actual (intended) amount to the user, while the tag will be sent a request for a higher amount. More seriously, such a malicious reader can also collude with a ghost and then succeed in purchasing an item much more costlier than what the user intended to buy [14]. As discussed in Section 1, addressing this reader-and-ghost relay attack requires transaction verification, i.e., validation that the tag is indeed authorizing the intended payment amount. Note that selective unlocking is ineffective for this purpose because the tag will anyway be unlocked in the presence of a valid (payment) context.

A display-equipped RFID tag can easily enable transaction verification for detecting reader-and-ghost attacks, as outlined in [40], [35], [14]. This, however, necessitates conscious user involvement because the amount displayed on the tag needs to be validated by the user and any user mistakes in this task may result in an attack. Distance bounding protocols have also been suggested as a countermeasure to the reader-and-ghost attacks [14]. However, these protocols are currently infeasible (as also reviewed in Section 6.2).

In this paper, we set out to explore the design of location-aware automated mechanisms for protecting against reader-and-ghost attacks. We note that under such attacks, the valid tag and the valid reader would usually not be in close proximity (e.g., the tag is at a restaurant, while the reader is at a jewelry shop [14]). This is in contrast to normal circumstances whereby the two entities would be at the same location, physically near to each other. Thus, a difference between the locations of the tag and the reader would imply the presence of such attacks. In other words, both the valid tag (credit card) and valid reader may transmit their locations to a centralized authority (issuer bank). This authority can then compare the information received from both entities and reject the transaction if the two mismatch.

We note that such a solution can be deployed, with minor changes on the side of the issuer bank, under the current payment infrastructure, where a card already shares a symmetric key with its issuer bank (as discussed in Section 2.2), and all communication takes place over secure channels. We only require that both the card and terminal measure their location information. Location information generated by both card and reader are then forwarded to the bank. The bank server decides whether to approve the transaction after comparing the location data received from the two ends. Fig. 3 illustrate the process of location-based proximity verification inside the current mobile payment infrastructure. The user-side card generates its location information $\text{loc}_{\text{card}}$ while the merchant-side reader generates its version of location information $\text{loc}_{\text{merchant}}$. The user-side card and the reader terminal then forwards its own location information $\text{loc}_{\text{merchant}}$ along with the card credentials to the bank for transaction verification and authorization. Since the integrity of $\text{loc}_{\text{card}}$ is protected (e.g., via MAC) with the key shared with the issuer bank before it is sent to the merchant’s terminal, which then forwards its own location information $\text{loc}_{\text{merchant}}$ along with the card credentials to the bank for transaction verification and authorization. Since the integrity of $\text{loc}_{\text{card}}$ is protected by the shared key between the card and bank, a malicious reader would be unable to change this value.

6 DESIGN AND IMPLEMENTATION

6.1 Location Sensing

Several positioning technologies can be used to get location information. The most popular positioning technologies to get location information include the satellite based-GPS, WiFi-based positioning system, and cellular network-based positioning system. Each of these positioning systems has its own favorable environment and performs much better than the others in terms of location estimation accuracy in most situations—hence a combination of them may not make sense to improve the overall accuracy [52].

GPS is generally used as the main source of location information and the major enabler for location-based services. It has world-wide availability and an accuracy of a few meters in location estimation—adequate enough for most civilian applications. However, the accuracy of GPS deteriorates inside buildings and in arrow urban canyons. Unlike GPS, WiFi positioning can provide good positioning results (with an accuracy of a few meters just like that of GPS) even indoors. However, it is prone to signal interferences and may not be always available due to the limited coverage of WiFi networks. Cellular network positioning is almost available both outside and indoors. However, it has poorer accuracy (50-100 meters) in location estimation.
Since location is used as a security control parameter in our approach, accuracy of location estimation can affect the security level we can achieve. For example, poor accuracy can cause a high false unlocking rate in selective unlocking and give more space for the adversary to cheat in proximity in server transaction verification. For this reason, the cellular network positioning technology is believed not a good candidate to use to get location information for security purpose.

In our experiments, we choose to use GPS to obtain location information for several reasons. First, although an NFC-enabled phone can have multiple communication interfaces including WiFi, it is unexpected a general RFID tag, even battery powered, will be equipped such an interface at the time being. Also, our proposed location-based security mechanisms are quite general and can work with a variety of positioning systems as long as they can provide adequate accuracy in location estimation. As low cost and inexpensive GPS sensors are available in the market (as discussed in Section 1), our purpose is to find out whether a low-cost commercial off-shelf GPS module with a few meter’s accuracy can meet the security requirement of the proposed schemes.

We also note, since in our server transaction verification, we use location information from both a reader and a tag to determine whether they are in proximity, we do not really need the physical location information (such latitude and longitude). Instead, location-specific information obtained by means of traditional ambient sensors can be used for proximity testing. This is based on the assumption that certain ambient information, extracted by the tag and reader at the same time (the time of transaction), will be highly correlated if the two devices are in close physical proximity. Therefore, if two sensors, one attached to the tag and the other to the reader, report mismatching ambient information, this will indicate that the tag and reader are (most likely) not at the same location or close to each other. We explored the idea of using ambient sensors for proximity check on NFC-equipped smartphones in a separate work [20].

In the following, we present more information about GPS background.

### 6.2 GPS Background

A GPS receiver derives its location by timing the signals sent by GPS satellites high above the Earth. The receiver uses the messages it receives from the satellites to determine the travel time of each message and computes the distance to respective satellite. These distances along with the satellites’ own locations are used with the possible aid of trilateration, to compute the position of the receiver.

GPS receivers can relay the gathered location data to a PC or other device using the NMEA 0183 specification [5]. This standard defines electrical signal requirements, data transmission protocol and time, and specific sentence formats for a 4800-baud serial data bus. Our approach is based on location and speed recognition. To obtain these two values properly, we need position, velocity, time (PVT) and data uncertainty (needed to establish the consistency of the data). GPGGA and GPRMC, the two most important NMEA sentences, are chosen for our implementation and experiments. GPGGA is an essential fix data that provide 3D location and accuracy (uncertainty) data. GPRMC has its own version of essential GPSPVT (position, velocity, time) data.

There are two methods to obtain the speed of the GPS unit. The first method calculates the speed indirectly from positional data differences and the corresponding time difference. The second method acquires the instantaneous Doppler speed directly from the GPRMC sentence. For our implementation, we use the Doppler speed because we can get this information instantaneously once we get a fix. Moreover, the Doppler speed is very accurate as it matches the readings from the car odometer in our experiments.

#### 6.3 Overview of WISP Tags

To evaluate the effectiveness and performance of the proposed location awareness techniques, we build proof-of-concept prototypes on the WISP tags. WISP's are passively powered RFID tags that are compliant with the Electronic Product Code (EPC) protocol. Specifically, we utilized the 4.1 version of the WISP hardware, which partially implements Class 1 Generation 2 of the EPC standard. These tags possess an onboard Texas Instruments MSP430F2132 microcontroller and sensors such as a three-axis accelerometer. The 16-bit MCU features an 8-MHz clock rate, 8 kilobytes of flash memory, and 512 bytes of RAM. WISP is chosen as our test platform because 1) it is the only existing programmable UHF RFID device, and 2) it has an extensible hardware architecture, which allows for integration of new sensors.

#### 6.4 System Overview

**GPS module.** As our test module, we have chosen the 66-channel LS20031 GPS receiver module from LOCOSYS Technologies in our experiments [2]. This module comes with an embedded ceramic patch antenna and GPS receiver circuits, which are designed for a broad spectrum OEM applications and outputs the data in more than six different NMEA GPS sentences to a TTL-level serial port. It provides us with a variable update rate of 1 to 5 Hz. This module also has a built-in micro battery for rapid satellite acquisition (which it does by preserving data). It also includes a LED indicator to indicate GPS fix or no fix [2].

In our experiments, we have configured the LS20031 to 1-Hz update rate, 57,600-bps serial communication rate and to output GGA and RMC NMEA sentences.

**Interfacing the GPS module with the WISP.** The LS20031 (GPS module) communicates via TTL level serial communication (UART), which is interfaced to the A channel communication port (used for UART, SPI, and I2C) on the WISP as shown in the block diagram above. The Rx communication on the LS20031 is only used for sending commands to configure it. The Tx port of LS20031 outputs the GPS NMEA sentences. Figs. 4 and 5 depict the block diagram as well as a picture from our experimental setup interfacing the WISP with our GPS module. As observed from Fig. 5, LS20031 has a small form factor and the WISP-LS20031 combination can be easily embedded within a traditional access card or toll card.

**Storing list of valid locations.** Since we have limited RAM, i.e., only 512 bytes on the WISP controller, we have to store
these valid location list on an external memory for the purpose of our selective unlocking mechanism (note that the transaction verification mechanism does not require the tag to store anything). Hence, we utilize the onboard EEPROM (8K) present on the WISP for storing the list of valid locations. Since this is an external memory to the controller (though onboard), one consideration we have to take into account is the time taken for the communication to take place between the controller and the EEPROM. This was found to be sufficiently small (about 3 ms) and feasible as we have the GPS output frequency of one sample per second.

We parse out location and the speed data from the GPS NMEA sentences. The latitude, longitude, and the speed are obtained from the GPRMC strings. The latitude and longitude data obtained are in degrees and the speed data is in knots. To avoid floating point numbers, the data are stored in the form of integers. To eliminate deviation in the GPS and errors, we average 10 such readings for 10 seconds and store these values. The lists of valid locations is then stored on the EEPROM and it serves as our reference to unlock the tag when the tag appears in one of the valid locations in the list.

The EEPROM is nonvolatile and so the list of valid location is retained unless it has to be changed or modified as per the requirements of the underlying application.

**Location sensing and computation.** For location sensing, we dynamically obtain the location data from the GPS continuously at the rate of 1 Hz, and compare it with the list of valid locations stored on the tag within a time span.

The issue of error tolerance plays a vital role in location recognition. To check whether an acquired location is a valid location in the location list, we test whether it falls within the square region centered at a valid location. The size of the square space depends on how much error tolerance we can afford. We conduct various experiments to find out the accuracies of location recognition based on different error tolerances. Since the values obtained from the GPS are in degrees, we map the degree error onto meters for easier understanding. We also have to consider the problem of different latitudes. Since the radii vary as we move across different latitudes, the error tolerance also varies. We found that for about 10 degrees of variation in latitude, the error tolerance varies by less than 1 meter, which is reasonably small and is feasible for most of the applications.

### 7 Experiments and Results

In this section, we present the experiments and associated results corresponding to our location-aware selective unlocking and transaction verification schemes.

#### 7.1 Selective Unlocking Experiments

We conducted three separate experiments to evaluate the performance of our selective unlocking mechanism based on location only, speed only, and both location and speed.

**Location tests.** In this experiment, we used location information as a selective control to lock/unlock the tag. We took the reading of five locations around the campus and stored them as valid locations where the tag should be in an unlocked state. We performed the test by driving around our university campus around these locations to measure the accuracy in recognition (20 recordings were taken for each error tolerance). A LED was used as an indicator for successful identification. This test was done using different error tolerances and our results were tabulated in Table 1. As an example, \((|x| \pm 2; |y| \pm 2)\) denotes an error tolerance of 2 meters centered around a valid location \((x, y)\) stored in the list. Referring to this table, we can conclude that we can successfully recognize valid locations under normal usage scenarios.

**Speed tests.** We make use of the instantaneous speed of the GPS receiver in our experiments. We found the instantaneous speed from the GPS receiver matches the reading of odometer in the car. We drove around the campus at different speeds (15, 25, and 35 mph) and five tests were conducted on each speed with each levels of error tolerance (results under the same error tolerance are clubbed together just to indicate the successful rate). When the speed falls within the predefined range, the LED on the WISP is turned on to indicate the tag was unlocked. Experiment results are shown in Table 2. We can conclude, referring to this table, that we can recognize the speed quite accurately.

<table>
<thead>
<tr>
<th>Test</th>
<th>Error Tolerance (meters)</th>
<th>% Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(</td>
<td>x</td>
</tr>
<tr>
<td>2</td>
<td>(</td>
<td>x</td>
</tr>
<tr>
<td>3</td>
<td>(</td>
<td>x</td>
</tr>
</tbody>
</table>
In this experiment, we used both location as well as speed as contextual parameters together to unlock the tag (as outlined in Section 4). This experiment is a combination of the previous two experiments. Here, the error tolerance for the location has to be set sufficiently high since the car is moving at a certain speed and the update rate of the GPS is 1 sample per second. Hence, we also have to consider the fact that the car moves a certain distance within that span of 1 second. For example, a car moving at 45 mph can travel around 20 meters in 1 second. So, an error tolerance of at least 20 meters has to be provided. This would not affect applications like car toll systems since most of the toll booths are located far away from other places, and, hence, the recognition area for the toll cards can be large [1]. In other words, using a higher error tolerance for such a system would not affect the system performance. As in prior experiments, an LED indicator was used for successful identification, which was later on used for unlocking the tag. The experimental results are shown in Tables 3 and 4, for two different speeds. We can observe that we were successfully able to unlock the tag based on the location and speed, and our accuracies improved considerably when the location error tolerance was increased.

### 7.2 Transaction Verification Experiments

We conducted another set of experiments for validating the effectiveness of our location-aware transaction verification scheme. The goal of these experiments was to determine the proximity (or lack thereof) between two devices—a valid tag and a valid reader—based on the location readings reported by their respective GPS receivers. In other words, we wanted to find out as to how accurately GPS sensing can be used to find out whether the two devices are in close proximity (e.g., at most 2 meters apart) or are far from each other (e.g., much more than 2 meters apart). Please recall that the former case represents a normal usage scenario for a typical payment token in which the user brings her card very close to the reader for processing a transaction. The latter, on the other hand, represents an attack scenario whereby the valid tag is at one location while the valid reader is at a different location [14].

We conducted two experiments to evaluate the proximity detection approach based on location data. By means of the first experiment, we wanted to determine the error tolerance of detecting proximity (within a distance of 2 meters). Note that when obtaining the location data from a GPS receiver at one particular location, we are subject to a maximum error around that point in a square region.

We connected a USB GPS sensor (GlobalSat BU-353) to the desktop which was in turn connected to our RFID reader, and set the distance between this receiver and the WISP receiver to be 2 meters. We then took 40 different samples from each of the two receivers simultaneously and from that we calculated the distance between the two receivers, and, thus, found out the range of maximum and minimum values. The minimum value was calculated to be 1.7821 meters and the maximum was 6.2093 meters. This means that even when the actual distance between the receivers is 2 meters, the distance reported by the GPS readings can vary between 1.7821 and 6.2093 meters. Therefore, a maximum error tolerance of 6.2093 meters could be used for the purpose of proximity detection.

Using the above error tolerance, we conducted our second experiment. Here, we wanted to determine the accuracy of proximity detection, based on the error tolerance of 6.2093 meters, when the distance between the two receivers was varied from 1 to 50 meters. The results of this experiment are reported in Table 5. As we can observe from this table, the accuracies corresponding to a distance of at most 2 meters are quite high as desired—this represents the normal use case (i.e., when no attacks occur). As the distance increases, the accuracies go down significantly, reaching a value of 0 percent for a distance of 20 meters or more. This means that if the adversary (illegitimate tag) is located more than 2 meters away from the valid tag, the possibility of the transaction being accepted are going to be low; in fact, the adversary does not stand a chance when he is located 20 meters or farther. This implies that if an adversary is at physically disparate location (e.g., at a jewelry store, while the valid tag is at a restaurant [14]), he will be easily detected and cannot succeed in the reader-and-ghost attack.
8 Discussion
In this section, we discuss related issues that may arise with respect to the proposed defenses in practice.

8.1 Preventing GPS Spoofing Attacks
Our location-aware defenses rely on the GPS infrastructure and, thus, may also be prone to the GPS associated vulnerabilities such as spoofing and jamming [59]. Successful spoofing experiments on standard receivers have been reported [44], [22], indicating commercial-off-the-shelf receivers do not detect such attacks. In the context of location-aware selective unlocking, the adversary can falsely unlock the tag if it can spoof the GPS signals coming from the satellites and feed in false location information to the GPS receiver (e.g., corresponding to a toll booth location even though the car/card is at a different location). Similarly, in the context of location-aware transaction verification, the adversary can, for example, fool the valid tag into thinking that it (the tag) is at a jewelry shop even though it is in a restaurant [14]. Commercial-off-the-shelf receivers do not detect such attacks.

Of existing GPS spoofing attack countermeasures [53], [37], [43], the one that is most suitable for the RFID setting is the scheme proposed in [43]. This scheme does not require any special hardware and does not rely on any cryptography. Instead, a GPS receiver in this scheme is augmented with inertial sensors (e.g., speedometers or accelerometers). The receiver can measure the discrepancy between its own predicted value (through inertial sensors) and measurements (through received GPS signals) to detect spoofing and replay attacks. The scheme is applicable to any mobile RFID tag setting, such as a toll card.

Since WISP already has an inertial (3-axis accelerometer) sensor onboard, we have the convenience of implementing the idea proposed in [43] against the GPS signal spoofing attack. The flowchart of our GPS detecting algorithm implementation is shown in Fig. 6. In our implementation, only two dimensions of the acceleration data have been taken into consideration because we are assuming that the tag is horizontally fixed on vehicle and the vehicle is always running in a horizontal plane. We compare the acceleration derived from the accelerometer data with the one derived using the speed provided by the GPS data over a short interval of time. When the difference between GPS calculated acceleration data and accelerometer data exceed a certain threshold, we consider the former as a possible spoofed data. We repeat this test and if spoofed data are being detected more than five times, we consider the tag to be under attack, and, thus, switch the tag into the locked state. To further reduce computation cost, we have used the square function for difference calculation instead of the square root function since square root is more computationally extensive for the WISP.

By adding inertial detection, we decrease the possibility of performing a successful signal spoofing attack thereby adding another layer of security to our system. However, this approach detects only the inertial abnormalities but not the location abnormalities. Thus, it only applies to situations where GPS receivers are mobile. Recently, a very interesting work on the requirements to successfully mount GPS spoofing attack has been reported [56]. The authors show that it is easy for an attacker to spoof any number of individual receivers. However, the attacker is restricted to only a few transmission locations when spoofing a group of receivers—even when they are stationary—while preserving their constellation (or mutual distances). Moreover, conducting spoofing attack on a group even becomes impossible if the group can hide the exact positioning of at least one GPS receiver from the attacker (e.g., by keeping it mobile on a vehicle) since in such cases the attacker cannot adapt to its position [56]. This suggests a cooperative detection scheme where multiple GPS receivers can work together to detect GPS spoofing attacks by also checking their mutual distances. Although it is still hard to foresee this countermeasure can be applied in current RFID application settings, it does state that a network of GPS receivers (or GPS-enabled devices) can be setup on the field to monitor GPS signals when it is necessary and when spoofing attack is a real menace.

8.2 GPS Initialization
A GPS can have either a cold start or hot start. The hot start occurs when the GPS device remembers its last calculated position and the satellites in view, the almanac (i.e., the information about all the satellites in the

![Fig. 6. Flowchart for detecting GPS spoofing attack.](image-url)
Delay due to GPS initialization, especially cold start, might be unbearable for delay-sensitive applications such as toll cards. However, in the toll card application, delay can be minimized by powering the tag with battery (which is the current power supply of most commercial toll cards) or the vehicle so that the GPS can always keep an updated view of the set of satellites with which it can get a fix immediately. In the building access card application, it is not reasonable to have an always-connected GPS receiver. However, since the receiver is powered up in the same place—e.g., office building entrance—as it was shut off last time under normal usage case, we can force the GPS receiver to do a hot start by remembering its last location (storing the location in nonvolatile storage). Moreover, the building access card application is more delay tolerant than the toll card application. That is, even the GPS receiver has to have a cold start, 35 seconds (time to have a cold start for the toll card application). That is, even the GPS receiver has to have a cold start, 35 seconds (time to have a cold start for the receiver we used in our experiments) might still be tolerable to most users.

8.3 Dealing with Failure Reading in RFID Toll Systems

Our speed-based unlocking scheme for toll cards only works when cars pass by the toll gates at the recommended speed. When a car actually does not pass toll gates at recommended speed, its toll card will be kept in locked state. The toll reader, hence, cannot read out the card information and the corresponding driver’s account, thus, cannot be successfully charged. So we need to deal with reading failure due to driver’s not driving at the recommended speed accidentally or intentionally. Actually, there already exists mechanism, which deals with failure reading in current RFID toll road system deployments. Current deployments rely on a combination of a camera, which takes a picture of the car, and a RFID reader, which searches for a drivers window/bumper mounted transponder to verify and collect payment. The system sends a notice and fine to cars (identified through either tag information or pictures taken by the camera) that pass through without having an active account or paying a toll. Our speed-based unlocking scheme can work together with the existing camera-based mechanism, and drivers are obligated to drive at the recommended speed to avoid fines.

9 Conclusion and Future Work

In this paper, we reported a new approach to defend against unauthorized reading and relay attacks in some RFID applications whereby location can be used as a valid context. We argued the feasibility of our approach in terms of both technical and economical aspects. Using location and derived speed information, we designed location-aware selective unlocking mechanisms and a location-aware transaction verification mechanism. For collecting this information, we made use of the GPS infrastructure. To demonstrate the feasibility of our location-aware defense mechanisms, we integrated a low-cost GPS receiver with a RFID tag (the Intel’s WISP) and conducted relevant experiments to acquire location and speed information from GPS readings. Our results show that it is possible to measure location and speed with high accuracies even on a constrained GPS-enabled platform, and that our location-aware defenses are quite useful in significantly raising the bar against the reader-and-leech attacks.

As an immediate avenue for further work, we intend to further optimize and fine-tune our location detection algorithms for better efficiency on resource-constrained RFID platforms and improved tolerance to errors whenever applicable. Additionally, we are exploring the use of ambient sensors to determine proximity based on location-specific sensor information for the second security primitive secure transaction verification. We will also evaluate the promising of proposed techniques by means of usability studies.

Acknowledgments

The work of Di Ma, Tuo Xiang, and Yan Zhu was partially supported by the US National Science Foundation (NSF) Grant CNS-1153573. The work of Nitesh Saxena was partially supported by the NSF Grant CNS-1201927.

References

Di Ma received the BEng degree from Xi’an Jiaotong University, China, the MEng degree from Nanyang Technological University, Singapore, and the PhD degree from the University of California, Irvine, in 2009. She is an assistant professor in the Computer and Information Science Department at the University of Michigan-Dearborn, where she leads the Security and Forensics Research Lab (SAFE). She was with IBM Almaden Research Center in 2008 and the Institute for Infocomm Research, Singapore in 2000-2005. She is a member of the IEEE.

Nitesh Saxena received the bachelor’s degree in mathematics and computing from the Indian Institute of Technology, Kharagpur, India, the MS degree in computer science from the University of California, Santa Barbara, and the PhD degree in information and computer science from the University of California, Irvine. He is an assistant professor in the Department of Computer and Information Sciences at the University of Alabama at Birmingham (UAB), and the founding director of the Security and Privacy in Emerging Systems (SPIES) group/lab. Before joining UAB, he was an assistant professor in the Department of Computer Science and Engineering at the Polytechnic Institute of New York University. He has also previously worked at Nokia Research Center, Finland, and at INRIA Rhone-Alpes, France. He is a member of the IEEE.

Yan Zhu is an associate professor in the Institute of Computer Science and Technology at Peking University and a visiting research scientist in the Computer and Information Science Department at the University of Michigan Dearborn, 2012. He was with the Department of Computer Science and Engineering, Arizona State University, as a visiting associate professor in 2008-2009.

For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.