Security from Location

Di Qiu, Dan Boneh, Sherman Lo and Per Enge
Stanford University
United States of America

1. Introduction

The emergence of the Internet and personal computers has led to an age of unprecedented information content and access. The proliferation of Internet connectivity, personal computers, and portable, high density data storage has put volumes of data are at one's fingertips. While the spread of such technology has increased efficiency and knowledge, it has also made information theft easier and more damaging.

The emerging problems have made the field of information security grow significantly in recent years. Geoencryption or location-based encryption is a means to enhance security. Precise location and time information can be used to restrict access of the system or equipment at certain locations and time frames (Qiu et al., 2007). The term "geo-security" or "location-based security" refer to the authentication algorithm that limits the access (decryption) of information content to specified locations and/or times. More generically, the restriction can be based on any set of location-dependent parameters. The algorithm does not replace any of the conventional cryptographic algorithms, but instead adds an additional layer of security.

When a device wishes to determine its position, it does two things (Qiu et al., 2010). First, the hardware uses an antenna and receiver to capture and record a location measurement. Second, the location measurement is converted into a global position in the form of longitude and latitude. Most often these two steps are conflated, and both are seen as necessary to enable location-based applications. In this paper we show that for many security applications only the first step is needed: there is no need to accurately map the location measurement to an accurate global position. Therefore, these location-based security applications can be implemented using a variety of radio frequency (RF) signals, including broadcast communication signals, such as AM/FM, cellular, DTV, Wi-Fi, etc, navigation signals, and an integration of various signals.

While GPS provides accurate position data, other location services are far less accurate. LOng RAnge Navigation (Loran), for example, uses a 3km wavelength, and standalone Loran has an absolute accuracy of several hundred meters (Loran-C, 1994). Loran-C, the most recent version of Loran in use, is a terrestrial navigation system originally designed for naval applications. Its modernized version, enhanced Loran (eLoran), together with differential corrections can achieve an accuracy of 8 to 20 meter. This paper uses standalone Loran-C, which has good repeatable accuracy but low absolute accuracy, as a case study and shows that high absolute accuracy is not a requirement for a number of location-based security applications. As with all radio-based systems, Loran-C radio signals are distorted by buildings and other objects causing measurements to change greatly over short distances. Our main result shows that one can exploit these chaotic changes to obtain a precise and reproducible geotag with an accuracy of about 20 meters. Reproducibility means that measurements at the same location at different times always produce the same tag. While there is no way to map location measurements to an accurate position, there are still many applications, primarily security applications, for which a reproducible and precise tag is sufficient.

We build a reproducible and precise tag using recent results from biometric authentication for location-based security applications. In particular, we rely on fuzzy extractors and secure sketches, originally designed for fingerprint-based authentication. The idea is to store some public information that enables anyone to convert an erroneous measurement into a consistent tag. We develop specific fuzzy extractors designed to handle radio-type errors. The challenge is to correct for signal variations due to day/night, humidity, and seasonal changes.

The rest of the chapter is organized as follows. Section 2 develops a standardized process to quantify the precision, reproducibility and security of a geotag for security applications. Section 3 provides definitions and background information on fuzzy extractors. The design and implementation of fuzzy extractors for location-based security discussed in Section 4 will apply to all radio-based signals. We use Loran-C as a convenient example and evaluate the geotag performance using real data, which will be addressed in Section 5.

2. Geo-security

2.1 System model

The geo-security system works in two steps, calibration and verification, as illustrated in Figure 1. The calibration phase builds the database of geotags for service areas: $\Im = \{T(\ell, t), \forall \ell \epsilon \mathcal{L}\}$, where T is the geotag of the calibration associated with location ℓ, and t represents the time interval when the geotag is generated. The use of time information for geotags is optional. The calibration phase requires one to survey the service areas with a location sensor, such as a Loran receiver that integrates a geotag generation module. Geotags associated with the calibrated areas are computed based on the recorded location information and stored on a database for future use. In the verification phase, a user derives a geotag $T'(\ell', t')\epsilon\Im,\ s.t.\ \ell'\epsilon\mathcal{L}$ using the same geotag generation device and matches it with the pre-computed ones in the database. If the two tags are matched, the user's location is validated and the authorization for an application is granted; otherwise, the authorization is denied.

2.1.1 Geotag generation

In this section we introduce two geotag generation methods: the deterministic approach and the binary approach. The methods differ in geotag representation, efficiency in computation and implementation in practice.

Let $x = f\left(s\left(\ell, t\right)\right)$, be the location-dependent parameters, where $s(\bullet)$ denotes the signals received at location ℓ and time t, and $f(\bullet)$ is the function performed in a receiver. Typical functions in a receiver include signal conditioning, digitizing, and parameter extraction. The extracted x is a vector $x = [x_1,\, x_2, \ldots,\, x_n]^T \epsilon \Re^{n\times 1}$, where n is the number of location-dependent parameters.

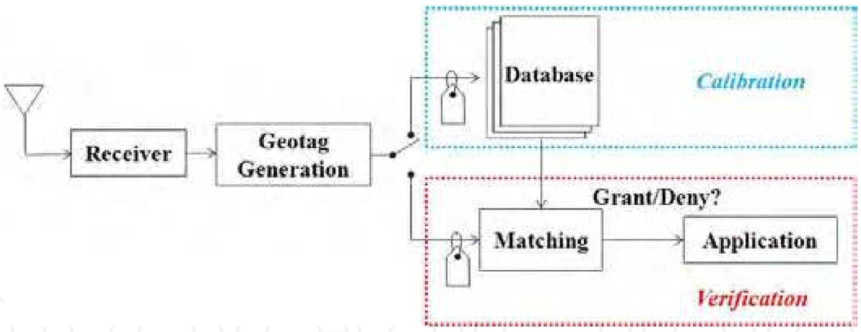


Fig. 1. Geo-security system: Calibration and verification phases

The deterministic approach simply takes the location-dependent parameter vector as a geotag, shown in Equation (1). This technique is similar to the location fingerprinting except that a geotag is computed from various location-dependent parameters rather than the received signal strength (Bahl & Padmanabhan, 2000).

$$T = \epsilon \Re^{n \times 1} \tag{1}$$

The binary geotag generation algorithm consists of three steps: a receiver function $f(\bullet)$ to extract location dependent parameters from the received signals $s(\ell, t)$, a quantizer $\mathcal{E}(\bullet)$ to quantize the parameters with adequate step sizes $\Delta(\ell)$, and a mapping function $\hbar(\bullet)$ to convert the quantized parameters into a binary string T. The binary mapping process can be done using a hash function, which is one-way and collision resistant. A one-way hash function is a fundamental building block in many cryptographic algorithms and protocols (Schneier, 1996), and outputs a fixed-length hash value regardless the length of inputs. One-way-ness means that it is easy to compute but hard or computationally infeasible to invert the function. In addition, since it is collision resistant, it is hard to generate the same hash values from two different inputs. Let q be the quantized parameter vector; its calculation is illustrated in Equation (2). All of these vectors x, q, and Δ have the size n. The quantization steps can be determined based on the standard deviations of the location dependent parameters to allow a certain degree of variations.

$$q_i = \mathcal{E}(x_i) = k; x \epsilon S_i = [k\Delta, (k+1)\Delta), k = 1, ..., N, \tag{2}$$

where S is the partition set and N indicates the number of quantization levels corresponding to a particular Δ. Thus the binary geotag can be calculated as

$$T = \hbar(q) \, \epsilon \, \mathbb{Z}^{m \times 1}. \tag{3}$$

2.1.2 Geotag matching

We next describe different matching algorithms for the two geotag generation functions. Two matching algorithms – the nearest neighbor method (NNM) and the probabilistic approach – can be applied to the deterministic geotag.

Let $\mathcal{M}$ denote the matching function. NNM is a common technique (Roos et al., 2002) used for indoor location estimation and pattern matching. The algorithm measures the distance between the location parameter vector from the verification phase T' and the previously stored vectors in the database, $\Im$. The generalized distance measure D is defined in Equation (4), where w is a weighting factor and p is the norm parameter. For instance, $w = 1$ and $p = 2$ represent the Euclidean distance. Based on the calculated distances between T' and the previously computed $T \epsilon \Im$, the geotag that gives the minimum distance is chosen. It is necessary to set an upper bound d_0 to guarantee that the location is registered at the calibration phase. A modification of NNM that uses the standard deviation σ of the location parameters is called the weighted nearest neighbor method (WNNM). The new distance measure is shown in Equation (5), where C is a covariance matrix, $C = E\{(x - \bar{x})^2\}$ and $\bar{x}$ is the mean value of location-dependent parameters. The matching function for the deterministic geotag is illustrated in Equation (6), where $\tilde{T}$ is the geotag associated with the authorized location.

$$D(x, x') = \frac{1}{n}(\textstyle\sum_{i=1}^{n} \frac{1}{w_i}|x'_i - x_i|^p)^{\frac{1}{p}} \tag{4}$$

$$D(x, x') = \left[(x - x')^T C^{-1}(x - x')\right]^{\frac{1}{2}} \tag{5}$$

$$\mathcal{M}(\tilde{T}, T') = \begin{cases} 1 & \text{if } \underset{T\epsilon\Im}{\arg\min} D(T, T') = \tilde{T},\ D(T, T') \leq d_0; \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

The probabilistic approach models a geotag with a conditional probability, and uses Bayesian methods to estimate the location (Roos et al., 2002). Both the location-dependent parameters and the standard deviations are estimated at the calibration phase. Assuming that the location-dependent parameters have Gaussian distributions, we use the probability density function shown in Equation (7) to compare the calculated likelihoods. The geotag that gives the maximum probability is chosen. The corresponding matching function is shown as follows:

$$P = \frac{1}{n}\textstyle\sum_{i=1}^{n}[\frac{1}{\sqrt{2\pi}\sigma_i} exp(-\frac{(x'_i - x_i)^2}{2\sigma_i^2})] \tag{7}$$

$$\mathcal{M}(\tilde{T}, T) = \begin{cases} 1 & \text{if } \underset{T\epsilon\Im}{\arg\max} P = \tilde{T}; \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

The matching process for a binary geotag only involves the correlations between T' and the previously stored ones. The correlation function is shown as follows:

$$\mathcal{M}(\tilde{T}, T') = \begin{cases} 1 & \text{if } \frac{1}{m}\sum_{i=1}^{m} \tilde{T}(i) \oplus T'(i) = 1,\ \forall \tilde{T} \epsilon \Im; \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

2.2 Loran-C for geo-security

The most important required feature of a signal for geo-security is its ability to generate a strong geotag. The strength of the geotag is determined by the quantity and quality of location-dependent signal parameters. By the quantity, we mean the number of different location-dependent parameters that can be generated. By the quality, we mean the amount of unique location-dependent information provided by each parameter. The information content is related to the spatial decorrelation of the parameter. Greater spatial decorrelation results in more unique information. By having many parameters each providing its unique information content, we can generate a strong geotag.

At the same time, it is desirable to have the parameters be relatively insensitive to temporal changes, which weaken the uniqueness of the information. Temporal variations essentially reduce the uniqueness of the location-dependent information. As a result, repeatability and repeatable accuracy are desirable qualities. They allow a user to have his location-dependent parameters or the derived geotag at one time—and still have those parameters valid at a later time. In other words, the signal characteristics should be consistent enough so that when the user is ready to authenticate, measurements at the same location will yield the same previously generated geotag. These are several features that are highly desirable.

In addition, the signal should have anti-spoofing capabilities. If the signal is vulnerable to spoofing, it may be possible for an attacker to bypass the location check and authenticate correctly. Furthermore, it is desirable that the signal be available indoors. This is because many of the anticipated applications of geo-security will likely occur indoors. This includes applications such as the management and distribution of secure digital data. Often, it is good if this data is only accessible inside certain buildings.

Loran-C is a terrestrial, low frequency, pulsed navigation system that operates in much of the northern hemisphere (Loran-C, 1994). Although the absolute accuracy of standalone Loran-C is not comparable to GPS, it has several advantages over GPS for security applications. First, Loran uses static transmitters and, as a result, its signals provide many parameters that are location-dependent. Each parameter offers different certain amount of information or potential information density. Parameters with higher information density result in stronger security. This is important, as the security strength of the geotag is derived from the information used to generate it. A combination of various parameters and the accuracy of these parameters increase the security strength. Second, Loran has good repeatable position accuracy, which benefits the design and guarantees the reproducibility of the geotag. Furthermore, Loran-C has good regional coverage in Northern Europe and much of East Asia like China, Japan, and Korea. Although the transmission of Loran-C signals in North America has been terminated in Feb. 2010, the decision with eLoran has yet to be made. eLoran will have a data channel (e-Loran, 2007). While some uses of the data have been defined, others have not. Therefore, several message types have been left unassigned to support useful application such as location-based security in the course of eLoran design. Loran antenna size may have been a practical issue in many applications. Recent research (Lee et al., 2009) has shown that a miniature H-field antenna of 2x2 cm can be achieved. With this size, a Loran H-field antenna can be easily fit into a number of portable electronic devices.

2.3 Applications

We discuss a number of potential security applications where the desired properties of geotags – high spatial decorrelation and reproducibility – come into play. Different geotag generation and system implementation methods should be applied to achieve optimized performance for various applications.

2.3.1 Digital manners policies (DMP)

Technologies for digital manners (DMP) (Hruska, 2008) attempt to enforce manners at public locations. A DMP-enabled cell phone can be programmed by the phone provider to turn off the camera while inside a hospital, a locker room, or a classified installation. Or the phone can be programmed to switch to vibrate mode while inside a movie theater. Many other applications have been considered. Although these ideas are highly controversial (Schneier, 2008), we only focus on the technical contents and feasible implementation of the ideas.

To implement DMP one assumes that the device needs to know its precise location. We argue that this is incorrect. Using our radio-based tag, one can build a list of geotags where the camera is to be turned off. The device downloads an updated list periodically. When the device encounters a geotag on this blocklist, it turns the camera off. When the device leaves the blocked location the camera is turned back on. Hence, digital manners are enforced without ever telling the device its precise location.

A DMP system must survive the following attack: the attacker owns the device and tries to make the device think it is somewhere else. Since most places are not blocked, any location confusion will do. To survive this threat any location-based DMP system must make the following two assumptions:

- First the device, including the antenna connection, must be tamper resistant. If the antenna connection is not protected then anyone can tamper with signals from the antenna. The simplest attack is to add a delay loop to the antenna. Since location measurements are time based, the delay loop will fool the device into thinking it is somewhere else.
- Second, it should be difficult to spoof the Loran-C radio signals by transmitting fake signals from a nearby transmitter. The safest defense against spoofing is cryptographic authentication for Loran-C signals. In our previous study we (Qiu et al., 2007) proposed a method for embedding TESLA (Perrig, 2002) authenticators into Loran-C signals to prevent spoofing. We point out that even without cryptography, spoofing Loran-C signals is far harder than spoofing GPS: In fact, GPS spoofers are commercially available and are regularly used by GPS vendors for testing their products.

Both assumptions are necessary to build an effective DMP system regardless of the navigation system used. Our goal is not to promote DMP but rather to show that an accurate DMP system can be built from standalone Loran-C signals.

2.3.2 Location-based access control

While DMP is a blocklisting application, access control is a whitelisting example. Consider a location-aware disk drive. The drive can be programmed to work only while safely in the data center. An attacker who steals the device will not be able to interact with it.

We consider two attack models:

- **Private locations:** suppose the device is located in a guarded data center and the attacker has no access to the insides of the data center. The attacker steals the device (say, while in transit (Sullivan, 2007)) and tries to make the device think it is still in the data center.

- **Public locations:** in this case the attacker has complete access to the data center and the attacker can measure the authorized geotag. After stealing the device the attacker can try to spoof the Loran-C signal to make the device think it is still in the data center. Unlike the DMP application where any location confusion was sufficient for the attacker, here the attacker must cause the device to think it is precisely in the right place in the data center, with 20 meter accuracy. Simply adding delay loops to the antenna will not work.

In both threat models we must assume that the device is tamper-resistant. Otherwise, the attacker can simply modify the device and bypass the location check. In the case of a public location we must also assume cryptographic authentication on Loran-C signals, as discussed in the DMP application.

Interestingly, for the private location settings, the unpredictability of the Loran-C geotag implies that we do not need any signal authentication nor do we need to protect the antenna connection to the device. In Section 5 we show that even if the attacker takes many measurements several hundreds of meters away (say in the parking lot) he still cannot tell for sure what tag to supply.

One option available to the attacker is to build a list of candidate geotags and try them one by one. In Section 5 we show that the list would need to include several dozen candidate tags. But the device can easily shutdown if it ever receives a sequence of incorrect geotags. Consequently, a trial and error attack will not get very far.

We note that location-based access control using encryption was studied by Scott and Denning (Scott & Denning, 2003) under the name Geoencryption, which uses physical locations, such as latitude, longitude and altitude measurements from GPS, for security applications. Our geotag derived from raw location measurements is more unpredictable and provides more information entropy.

3. Background on fuzzy extractors

In the previous section we showed applications for a precise and reproducible geotag. We now show how to build such tags using standalone Loran-C system. To ensure that our tags are reproducible we will make use of fuzzy extractors (Juels & Wattenberg, 1999; Dodis et al., 2004). Fuzzy extractors were originally designed for biometric authentication systems. Since biometric scanners introduce errors, one needs same way to extract a reproducible tag from the scanner's output. While biometric fuzzy extractors are designed with a specific error model in mind, here we need a fuzzy extractor tailored for the Loran error model.

3.1 Fuzzy extractors: Definitions

We follow the definitions in (Dodis et al., 2004). Measurements live in a set M which is equipped with a distance function denoted dis. Roughly speaking, $dis(x, y)$ is small if x is "close" to y.

Fuzzy extractor. A fuzzy extractor works in two steps. During the registration step one runs algorithm Gen on input $x \epsilon M$ to generate a public value P and a tag T. Later, given a noisy version of x, denoted x', one runs algorithm Rep on input x' and P to reproduce the tag T.

The idea is that if x and x' are fingerprint scans of the same finger, then x is "close" to x' and both should produce the same tag T. If T has sufficient entropy then it can used as a login password. Clearly we require that P reveal little or no information about the tag T.

Definition 1. A fuzzy extractor is a tuple (M, t_0, t_1, Gen, Rep), where M is the metric space with a distance function dis, Gen is a generate procedure and Rep is a reproduce procedure, which has the following properties:

If $Gen(x)$ outputs (T, P), then $Rep(x, P) = T$, whenever $dis(x, x') \leq t_0$. If $dis(x, x') \geq t_0$, then there is no guarantee T will be output. In addition, if $dis(x, x') \geq t_1$, $Rep(x', P) = T'$, and $T' \neq T$.

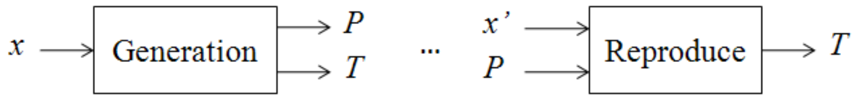


Fig. 2. Fuzzy extractor in action

3.2 Known constructions for fuzzy extractors

Initial constructions were proposed by Juels and Wattenberg (Juels & Wattenberg, 1999). Their scheme uses an error correcting code to handle the hamming metric on binary data. Juels and Sudan (Juels & Sudan, 2002) provide a fuzzy extractor for the set difference metric, which is the first construction for a non-hamming metric. Dodis (Dodis et al., 2004) gives precise definitions for the problem and provide constructions for hamming distance, set distance and edit distance.

All these schemes primarily apply to binary data which does not fit our settings where location measurements are vectors of real numbers. One exception is a construction of Chang and Li (Chang & Li, 2005) that can be adapted to give a fuzzy extractor for the scenario where one of the Loran-C transmitters is offline (e.g. for maintenance).

4. Generating a reproducible and precise geotag from Loran-C

Our goal is to build a reproducible and precise geotag from standalone Loran-C measurements. We first explain what a Loran-C measurement looks like and then discuss the error model for these measurements. Finally, we present a simple fuzzy extractor for this error model.

Loran-C measurements. Radio-based navigation uses signals from multiple transmitters to estimate the receiver's positions. Four transmitters on the west coast of the US, called the west coast Loran chain (GRI9940) are used for navigation in the western US. These four stations are located at Fallon, NV; George, WA; Middletown, CA; and Searchlight, NV. Pulses from this chain are broadcast every 0.0994 seconds (Loran-C, 1994). Fallon is the master station and the remaining three follow in sync. From each station we obtain three values, called location parameters or **features**, per pulse:

- Time-of-arrival (TOA) or time difference (TD): measures the propagation time from the transmitter to the receiver,
- envelope-to-cycle difference (ECD): measures carrier propagation rate, and
- signal-to-noise ratio (SNR).

An example measurement from the Middletown, CA station taken at Stanford is a triple: (496.8 microseconds, -0.145 microseconds, 41dB).

The exact meaning of these numbers is not important for our discussion here. What is important is that each transmitter produces a triple of real numbers (features) per pulse. Collecting the signals from all four stations gives a 12-dimensional real vector from which we wish to derive a geotag.

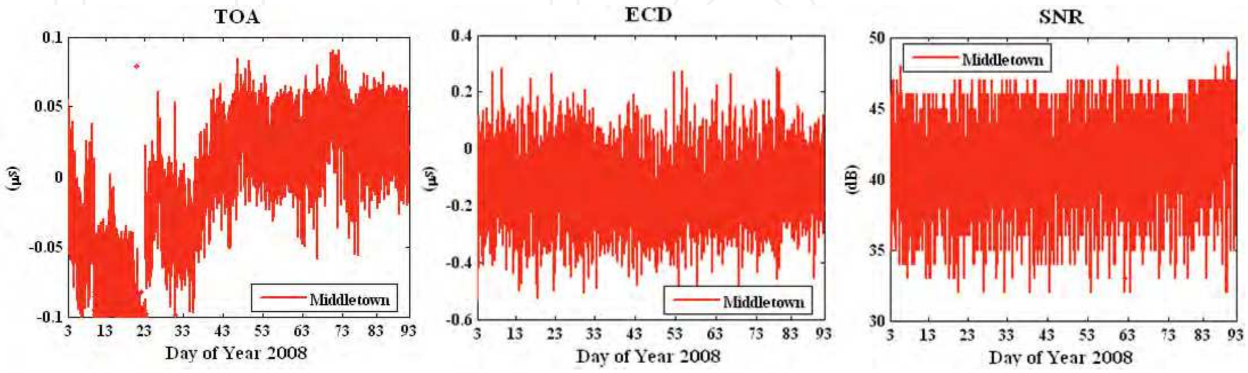


Fig. 3. Stanford seasonal monitor data for 90-day period for Middletown: (a) TOA; (b) ECD; (c) SNR.

Loran-C error patterns. Due to measurement errors and environmental changes, taking multiple measurements at the same location, but at different times, produces different 12 dimensional vectors. Figure 3 shows temporal variations in the triple (TOA, ECD and SNR) as measured from the Middletown station over a 90 day period. These measurements were taken at Stanford, CA. The wild swings in TOA, for example, reflect seasonal variations between winter and spring. We next explain the reason for these variations and how to model them.

- The most common error source is the thermal noise in all electronic devices, considered as white Gaussian noise. This noise cannot be eliminated and is always presenting in all electronic devices and transmission media.
- Many environmental factors cause signal variation, including temperature changes between night and day, changes in soil conductivity over time, humidity, local weather, etc. (Swaszek et al., 2007). In particular, temperature and humidity variations have a considerable effect on propagation speed. The extra delay in propagation time or TOA can introduce a position error of hundreds of meters (Lo et al., 2008). This particular error source in Loran is called additional secondary factor (ASF) and represents one of the largest error sources in Loran.
- Location vectors are continuous and need to be quantized. Quantization error, which is the difference between value of continuous feature and the quantized value, can lead to errors in the derived geotag. The quantization error is usually correlated with the two types of errors discussed above.
- The last type error results from maintenance of any radio-based system. A transmitter can go offline, in which case we lose all measurements associated with that station. Ideally, we would like this to have no effect on the geotag produced by our system.

A fuzzy extractor for Loran signals must take seasonal variations into account and can correct errors differently depending on the time of year.

4.1 Construction 1: Fuzzy extractor for Euclidean distance

We propose a fuzzy extractor when all Loran-C transmitters are present (Qiu et al., 2010). Thus the features are real numbers over R and Euclidean distance is sufficient for the distance metric. Let x be a location feature vector at registration while x' be the feature vector at verification time, Δ is the step size to quantize the feature. The distance $dis(x, x')$ can be bounded by adequate threshold. This threshold, δ, can be a design parameter. We need to develop a fuzzy extractor that can reproduce geotag T when the errors $|x - x'| \leq \delta$. The fuzzy extractor is designed to tolerate the random noise, biases and quantization errors.

Let the metric space $M = [A_i, B_i]^n, n = 12$ if we use the triple from four Loran-C stations. Thus x, x' and Δ are vectors that have n dimensions. The quantization step Δ is a design parameter and chosen by a user. We consider the distance measure for Loran-C features is L_∞ norm to be conservative.

$$dis(x, x') = \left(\max_i \frac{|x_i - x'_i|}{\Delta_i}\right)_{i=1}^{n} \tag{10}$$

The construction of fuzzy extractor for Euclidean distance is as follows: during calibration or registration, feature vector x is quantized to get T and store public value P, whereas, during verification, given a slightly different location feature x' and P, compute T'. P, T and T' are also n-dimensional vectors. P_i represents the i^{th} feature in vector P. The elements in vector T are integers but they are not necessarily positive. For instance, it is possible to result in a negative TD if the distance between the secondary station and a user is shorter than the distance between master station and the user. The basic idea of this fuzzy extractor is to adjust the offsets between the continuous features and the discrete ones due to quantization.

$$Gen(x) = \begin{cases} T = \lfloor \frac{x_i}{\Delta_i} \rfloor_{i=1}^{n} \\ P = \left(x_i - \Delta_i \lfloor \frac{x_i}{\Delta_i} \rfloor\right)_{i=1}^{n} \end{cases} \tag{11}$$

$$Rep(x', P) = \lfloor \frac{x'_i - P_i + \frac{\Delta_i}{2}}{\Delta_i} \rfloor_{i=1}^{n} = T' \tag{12}$$

Claim 1. If $dis(x, x') < \frac{1}{2}$, then a geotag T can be reproduced, that is, $T' = T$. This claim defines the reproducibility of geotags. If x' is measured at the same location of x, we can reproduce T when the distance of x and x' is less than $\frac{\Delta}{2}$.

Claim 2. If $dis(x, x') \geq t_1$, then a geotag $T' \neq T$. This claim defines the precision of geotags. If x' is measured at a different location but close to the location of x, it is not expected that x' achieves the same tag as x.

It is easy to see that our construction is a fuzzy extractor (as in **Definition 1**).

4.2 Construction 2: Secret sharing based fuzzy extractor for hamming distance

The distance metric in this construction is Hamming. The input to the fuzzy extractor is quantized feature vector q_x instead of x, where $q_x = \lfloor \frac{x_i}{\Delta_i} \rfloor_{i=1}^{n}$ is n-dimensional. The scheme is based on the property of secret sharing: a secret can be reconstructed given a subset of shared information. The construction is as follows:

- Create a polynomial $f(x)$, such that $f(i) = q_{x_i}, \forall i = 1, 2, ..., n$.
- Let m be an integer and $m < n$.
- $Gen(x) = \begin{cases} T = \langle f(1), f(2), ..., f(m) \rangle \\ P = \langle f(j), ..., f(j+n-m-1) \rangle \end{cases}, where\ j, ..., j+n-m-1 \notin \{1, ...n\}$.
- $Rep(x', P) = \begin{cases} f'(x) \\ T' = \langle f'(1), f'(2), ..., f'(m) \rangle \end{cases}$.

Claim 3. If $dis(q_x, q_{x'}) \leq n - m$, then a geotag T can be reproduced. When the hamming distance between two vectors is less than $n-m$, the polynomial $f(x)$ can be reconstructed with the assistance of P thus $T' = T$.

Claim 4. If $dis(q_x, q_{x'}) > n - m$, then a geotag $T' \neq T$. The precision of a geotag T relies on the features $x_1, ..., x_m$.

This construction increases reproducibility but reduces entropy because we only use m out of n features to compute a geotag.

5. Experimental results

In this section we use real standalone Loran-C data to evaluate the precision and reproducibility of Loran-C geotag and evaluate the effect of the Euclidean metric fuzzy extractor. We performed two experiments: (1) collected data at various test locations to examine the precision of geotags, and (2) collected data at one location over 90-day period to study the reproducibility of geotags.

5.1 Data at different locations evaluating tag precision

We selected three different environments, where our proposed location-based security applications may occur, to perform the precision test: parking structure, soccer field and office building. At each location we used multiple test points for five minutes at each test point. An H-field antenna and Locus Satmate receiver, shown in Figure 4, were used for the data collection. The receiver averages and outputs Loran location features every minute.

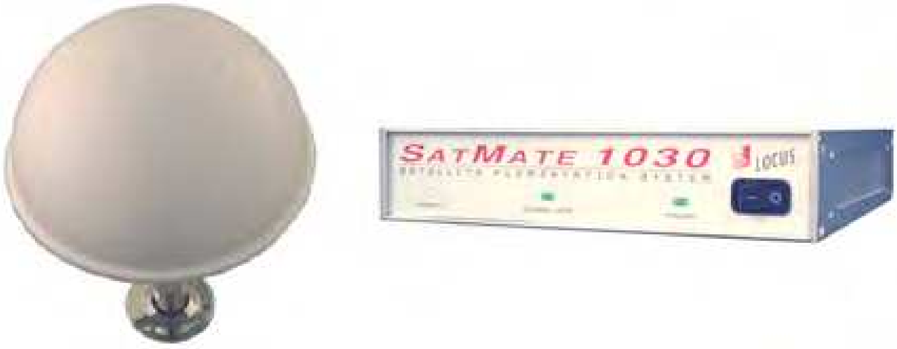


Fig. 4. Loran-C H-field antenna(left) and SatMate receiver (right)

- **Scenario 1.** The first data set was collected at 21 different test points on the top floor of a parking structure at Stanford University. This place has open sky view and no obstruction from the environments but there are some metal structures nearby. The altitude is relatively high compared with the other two scenarios. The dimension of the parking structure is approximately 70 x 50 meters.
- **Scenario 2.** The second data set selected 16 test points in a soccer field. This environment has some obstructions from trees and buildings. The field has a dimension of 176 x 70 meters so the distribution of the test locations are less dense compared to the other two scenarios.
- **Scenario 3.** The third data set, which includes 21 test points, was collected on the top floor both inside and outside a building. The concrete building with metal frames attenuates signal strength more but introduces more uniqueness in the location features, which can be beneficial to the computation of geotags.

We used the triple (TD, ECD, SNR) from four stations in the west coast chain (GRI 9940). Quantization steps are chosen based on the measured SNR. Low SNR signals are often attenuated more and pick up more noise. In general, features from low SNR stations are less consistent; thus larger quantization steps should be applied. We then created two-dimensional cells using Voronoi diagrams and mapped the tags into the cells accordingly. The color map is superimposed on the Google map. A color bar is used to label the hexadecimals of the first 16-bit of tag. This distribution plot can help us visualize how geotag varies in a two-dimensional view. Each black dot together with the numbered label at the center of the cells represents a test location.

The left of Figure 4 is the tag plot on the top floor of the parking structure, the middle plot represents the results of a soccer field, and the right plot shows the top floor/roof of Durand building. Loran signals are very sensitive to the environment, especially to metal structures. The re-radiation of signals from metals can cause more distortion to the RF signals thus higher precision or spatial variation of tags at certain locations. We observe this from the geotag maps of scenario 1 and scenario 3. The locations with very small separations still result in different geotags. It is worth to mention that only two stations, Fallon and Middletown, are used to compute tags for scenario 3 while the other two scenarios use all four stations from GRI 9940. Due to the low signal strength indoors, the SatMate receiver was not able to acquire the other two low SNR stations, George and Searchlight. The averaged precision of three different scenarios is as follows:

- The precision of Loran-C tags in the parking structure ranges from 8 meters to 35 meters. There are four locations that resulted in the same tag shown in dark blue on the left of Figure 5.
- The precision of tags in the soccer field is lower compared with that of the parking structure due to the large separations between the selected test locations or insufficient number of test points used. The averaged size of the colored cells that represents geotag is approximately 30 x 50 meters.
- Although the indoor signals are not good enough to solve a position fix because low-SNR signals are not able to track. The generation of a geotag does not rely on the solved position fix as the geotags are derived from location-dependent features. As a result, it is not required to have more than four transmitters to implement location-based security although more transmitters would provide more information entropy or longer

tag to the system. The smallest colored cell or the highest tag precision in this indoor scenario is approximately 5 meters depicted in purple in the middle of the right plot in Figure 4. An upper bound on actual tag precision at this location is the largest cell, 8 x 20 meters.

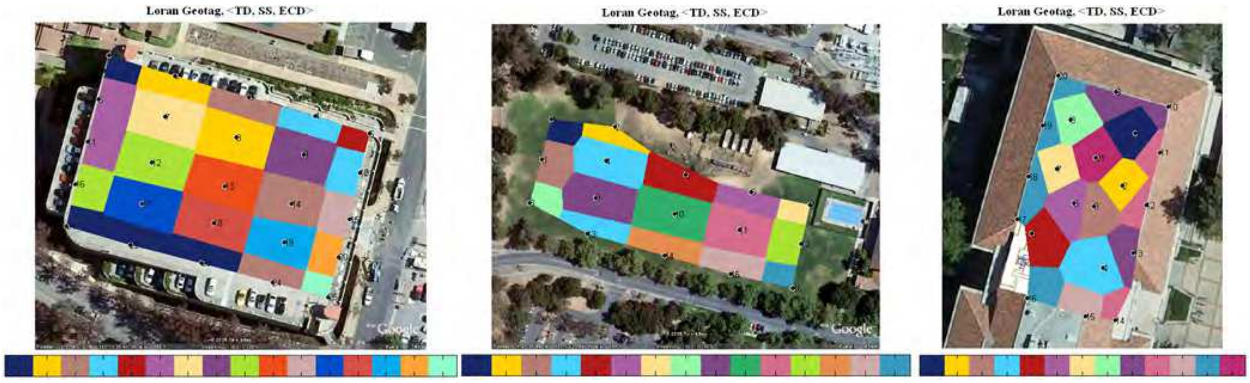

Loran Geotag, <TD, SS, ECD>
Loran Geotag, <TD, SS, ECD>
Loran Geotag, <TD, SS, ECD>

Fig. 5. Visualization of Loran geotags: (a) parking structure (left); (b) soccer field (middle); (c) Durand building (right)

5.2 Data at one location evaluating reproducibility

In this section we use the seasonal data shown in Figure 3 to compare the reproducibility of a geotag with and without a fuzzy extractor. Again same triple is used in this experiment. We use TD instead of TOA to minimize the impact of ASF errors: TOA of the master station is used as a reference to mitigate the temporal variations of secondary stations. Our experiments show that the standard deviation of TOA from Middletown is 12.19 meters and the standard deviation of TD from Middletown is reduced to 3.83 meters (Qiu et al., 2008). However, TD provides less information entropy in comparison with TOA as we lose the TOA entropy from master station.

Performance metrics. Before we discuss the experimental results from the seasonal data we introduce the performance metrics that help to quantify and measure the reproducibility of a geotag. The problem of deciding whether the derived geotag is authentic or not, can be seen as a hypothesis testing problem. The task is to decide which of the two hypotheses H_0 (accepting as an authorized user) or H_1 (rejecting as an attacker) is true for the observed location measurements. Location-based system makes two types of errors: 1) mistaking the measurements or derived tag from the same location to be from two different locations and accepting hypothesis H_1 when H_0 is true, called false reject; and 2) mistaking the measurements or derived tags from two different locations to be from the same location and accepting H_0 when H_1 is true, called false accept. Both false reject rate (FRR) and false accept rate (FAR) depend on the accuracy of equipments used, step sizes chosen to quantize location features and environmental conditions. These two types of errors can be traded off against each other by varying the quantization steps. A more secure system aims for low FARs at the expense of high FRRs, while a more convenient system aims for low FRRs at the expense of high FARs. Figure 6 illustrates the two error rates of geotags with the assumption that the probability distributions are Gaussian, which is not necessarily true in practice. The grey tails represent the false reject of an authorized user while the red area is the false accept of an attacker.

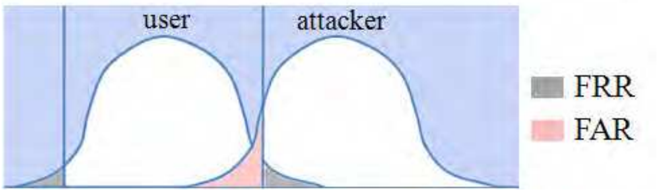


Fig. 6. Performance metrics illustration

Choosing a reliable quantization step for a location feature. Users' false reject rate significantly depends on the standard deviation of the features. Large standard deviation implies high temporal variations; thus the distance between the received features at verification and the ones at registration might be large. Therefore, the quantization step should be chosen to be proportional to the standard deviation σ of features.

In this analysis we show that the quantization step has to be larger than 4σ to achieve reasonably small FRR, less than 0.1. The FRR analysis is illustrated in Figure 7. The quantization step ranges from σ to 6σ. The x-axis is the feature offset between registration and verification. The y-axis is the estimated FRR. The solid lines are analytical results and we assumed the distribution of location feature is near-Gaussian after the ASF mitigation. The dots are derived using the seasonal data. We used ECD from four stations in this experiment. To estimate FRR we take the first day of the 90-day ECD data as registration to compute a geotag and the data from the rest of 89 days for verification. The experimental FRR is the number of days, in which the tags are matched with the registered tag on day one, divided by 89. The experimental results match well with the analytical curves. As expected, FRR increases as offset goes up and quantization step goes down.

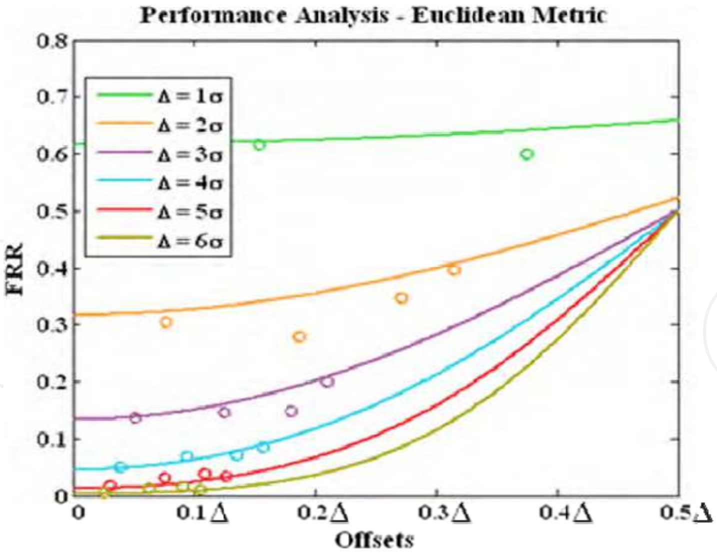


Fig. 7. FRR of a location feature

Using multiple features. The derived FRR in Figure 6 only represents the error rate of one particular location feature. Practically, multiple features are used to achieve more entropy, precision and higher difficulty in predicting the desired tag. However, one drawback using multiple features is that the FRR of the system is increased or reproducibility is reduced.

The system FRR can be estimated as $\Pi_{i=1}^{n} p_i$ if we assume the location features are independent from each other, where p_i is the error rate of one feature. Practically, location features are slightly correlated in some environments. For instance, the signal strength is inversely proportional to the propagation distance, which is determined by TOA. This is true when the antenna is placed in an open sky area and has no obstructions from surroundings. To solve the reliability problem using multiple features, secret sharing based fuzzy extractor can be used together with the Euclidean metric fuzzy extractor. Only a subset of features is used to compute tags thus the total FRR is limited.

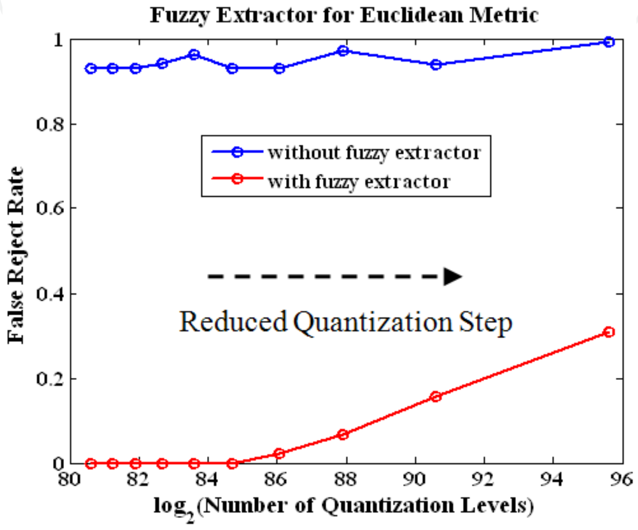


Fig. 8. Performance of Euclidean metric fuzzy extractor

Euclidean metric fuzzy extractor performance of multiple features. Now we use the triple from four stations to evaluate experimentally the performance of Euclidean metric fuzzy extractor. We reduce the quantization steps of the features gradually to observe the change of FRR and the number of quantization levels, which determine the entropy of geotag. The plot is shown in Figure 8. The blue line represents the FRR without the use of the fuzzy extractor while the red line is the results using the fuzzy extractor. As expected, the FRR is dramatically reduced after the use of the fuzzy extractor. The fuzzy extractor guarantees the measurements lying in the center of quantization interval. The graph shows that we can achieve total entropy of 86 bits with FRR is less 0.1 with adequate quantization steps.

5.3 Loran-C geotags are unpredictable

Next we ask whether Loran-C geotags are predictable from a distance. In this chapter unpredictability refers to the difficulty of an individual in predicting the Loran measurements at a given time and place. The temporal variations due to propagation path delay variations and skywave as well as the unexpected distortions in the RF signals due to local features such as buildings and large metallic structures can introduce randomness and entropy in the generation of a geotag, which makes attackers to take more time and effort to break into the system.

We discussed applications for this unpredictability test in Section 2.3. To justify the claim that Loran-C geotags are unpredictable, we perform two experiments.

While we cannot prove the difficulty of prediction mathematically as it is not possible to come up a universal model that suits for all the environments; however, we can show the nonlinear of the Loran-C features experimentally. The predictions can be based on path propagation, reflection, diffraction, diffuse wall scattering and transmission through various materials. The sum of all the components is taken to get TD, ECD and SNR. Moving objects like people can cause not only attenuation but also fluctuation. The irregularities make the prediction even harder.

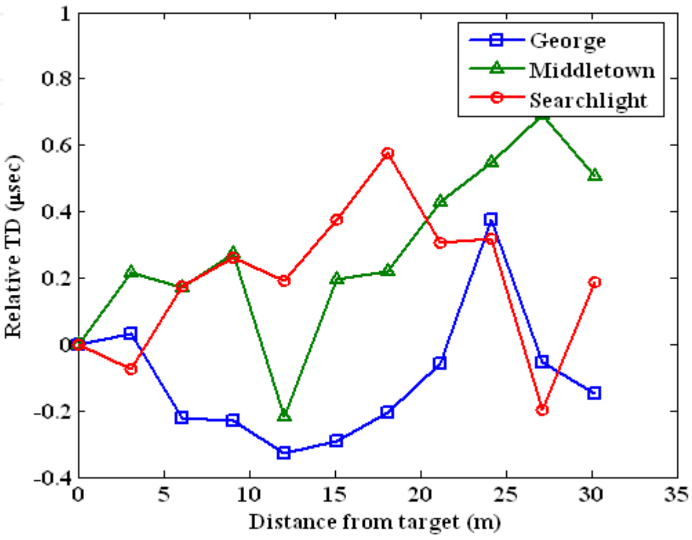


Fig. 9. Spatial variation of TD measurements collected in a parking structure

We perform the following two experiments to test the difficulty to predict a geotag. The first experiment uses the data set collected in a parking structure from 11 test points. The test locations are lined up in one dimension and the separation between adjacent points is approximately three meters. We chose the first point as our target or user location. Figure 8 plots the spatial variations of TD of George, Middletown and Searchlight. The x-axis is the measured distance of test points from the target point. The y-axis is the relative TD in microseconds. We zeroed out the means of the TDs to achieve the same scale for the measurements from three stations. The nonlinearity of the Loran-C measurements is clear from the graph. Low-SNR stations, George and Searchlight, are attenuated more from the obstructions in the environment compared to the strongest station Middletown. This results in more nonlinear variations in the low-SNR stations.

The second experiment uses the same data set collected in Durand building for the precision test discussed in Section 5.1. We chose the center point as our target point and measured Loran-C features with increasing distances from the target point. The point is shown as white dots in the plots of Figure 10. The color contour plot is again superimposed on the Google map. The color bar shown at the bottom represents feature values of various locations. Figure 10 illustrates the spatial variations of TD, ECD and Signal strength measured from Middletown. If feature variations are linearly proportional to distance, the color of the map should change from blue to red gradually with equal diameter. We observe that ECD are more nonlinear in comparison with TD and signal strength because phase is very sensitive to building structures and environments. The non-linearity of location features can significantly benefit the design of location-based security applications as it results in the features are highly unpredictable.

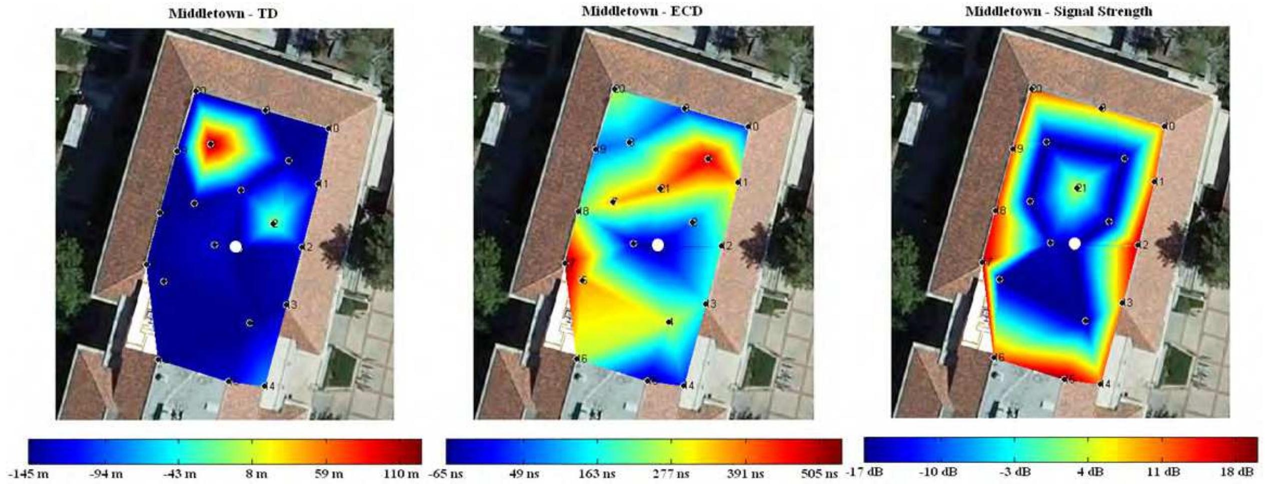


Fig. 10. Spatial variation of location data from Middletown in Durand building: (a) TD; (b) ECD; (c) Signal strength.

6. Conclusion

We showed that a radio navigation system with high absolute accuracy and low repeatable accuracy such as standalone Loran-C can be used to generate a precise and reproducible geotag. A geotag is computed from location-dependent features and can be used for a number of security applications. A geotag is not a replacement but builds on the conventional security schemes. We discussed applications to DMP, inventory control and data access control.

Fuzzy extractors were developed for radio-based signals to achieve high consistency. Euclidean metric fuzzy extractor and Hamming metric fuzzy extractor were designed for different location measurement errors. Adequate quantization step should be chosen as it determines the system performance. FAR and FRR can be traded off by varying the quantization steps of location features. We used Loran-C real data to show that the Euclidean metric fuzzy extractor significantly improves the reproducibility of a generated geotag. In addition we proved that the Loran-C location features can achieve high spatial variation using measurements at three different sites, a parking structure, a soccer field and an office building. In addition, we gave evidence that a geotag is unpredictable from a distance, which is beneficial to location-based security applications.

This paper only focused on the evaluation of geo-security using Loran-C as a case study; however, there are many available radio signals that might be feasible to implement geo-security, such as digital television, cellullar, Wi-Fi, and RFID. The proposed location-based security technique needs to be validated and compared with case studies. Future work shall be directed toward design of experimental setups, evaluating the feasibility and performance of each signal, comparing the different signals in terms of performance, usability and cost, and serivce coverage.

7. References

Enhanced Loran (eLoran) Definitions Document (2007). International Loran Association. URL: *http://www.loran.org/ILAArchive*

Loran-C Signal Specifications (1994). United States Coast Guard (USCG), COMDTINST M15662.4A, May 1994.

Bahl, P. & Padmanabhan V.N. (2000). RADAR: an in-building RF-based user location and tracking system, *Proceedings of IEEE in INFOCOM 2000*, IEEE, Vol. 2 (2000), pp. 775-784.

Boyen, X. (2004). Reusable cryptographic fuzzy extractors, *Proceeding of the 11th ACM Conference on Computer and Communications Security*, ACM Press, pp. 82-91.

Chang, E. & Li, L. (2005). Small secure sketch for point-set difference, *Cryptology ePrint Archive, Report 2005/145.*

Dodis, Y.; Reyzin, L. & Smith, A. (2004). Fuzzy extractors: How to generate strong keys from biometrics and other noisy data, *Eurocrpt'04*, Springer-Verlag, Vol. 3027 of LNCS, pp. 523-540.

Hruska, J. (2008). Microsoft patent brings miss manners into the digital age, *Arstechnica Hardware news*, June 11, 2008.

Juels, A. & Sudan, M. (2002). A fuzzy vault scheme, *Proceeding of IEEE Intl. Symp. on Information Theory*, IEEE Press, pp.408, Lausanne, Switzerland.

Juels, A. & Wattenberg, M. (1999). A fuzzy commitment scheme, *Sixth ACM Conference on Computer and Communications Security*, ACM Press, pp.28-36, 1999.

Lee, D.; Best, S.; Hanna, D. & Rosario, E. (2009). A miniature Loran H-field antenna for low-profile conformal hybrid applications, *Proceeding of ION ITM 2009*, Institute of Navigation, Jan. 2009, Anaheim, California, United States.

Lo, S.; Wenzel, R.; Johnson, G. & Enge, P. (2008). Assessment of the methodology for bounding Loran temporal ASF for aviation, *Proceeding of ION NTM 2008*, Institute of Navigation, Jan. 28-30, 2008, San Diego, California, United States.

Perrig, A.; Canetti, R.; Tygar, J.D. & Song, D. (2002). The TESLA broadcast authentication protocol, *CryptoBytes*, 5:2, Summer/Fall 2002, pp. 2-13.

Qiu, D.; Boneh, D.; Lo, S.; Enge, P. Reliable location-based srvices from radio navigation systems, *Sensors* 2010, *10*, 11369-11389.

Qiu, D.; Lo, S.; Enge, P.; Boneh, D. & Peterson, B. (2007). Geoencryption using Loran, *Proceeding of ION NTM 2007*, Institute of Navigation, Sep. 25-28, 2007, San Diego, California, United States.

Qiu, D.; Lo, S. & Enge, P. (2008). A measure of Loran location information, *Proceeding of IEEE/ION PLANS 2008*, Institute of Navigation, May 6-8, 2008, Monterey, California, United States.

Roos, T.; Myllymaki, P.; Tirri, H.; Misikangas, P. & Sievanen, J. (2002). A probabilistic appraoch to WLAN user location estimation, *International Journal of Wireless Information Networks*, 9(3): 155-164, July 2002.

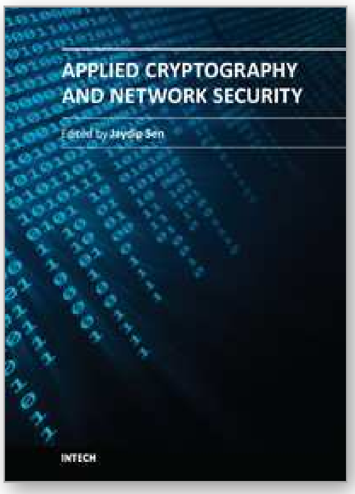
Schneier, B. (1996). *Applied Cryptography*, John Wiley & Sons, ISBN 0-471-11709-9.

Schneier, B. (2008). Kill switches and remote control, *A blog covering security and security technology*, July 1, 2008.

Scott, L. and Denning, D. (2003). A location based encryption technique and some of its applications, *Proceedings of ION NTM 2003*, Institute of Navigation, Jan. 22-24, 2003, Anaheim, California, United States.

Sullivan, B. (2007). The biggest data disaster ever, *MSNBC news.* Nov. 30th, 2007.

Swaszek, P.; Johnson, G.; Hartnett, R. & Lo, S. (2007). An investigation into the temporal correlation at the ASF monitor sites, *Proceedings of ILA 36th Annual Meeting 2007*, International Loran Association, Oct. 14-17, 2007, Orlando, Florida, United States.



Applied Cryptography and Network Security
Edited by Dr. Jaydip Sen

ISBN 978-953-51-0218-2
Hard cover, 376 pages
Publisher InTech
Published online 14, March, 2012
Published in print edition March, 2012

Cryptography will continue to play important roles in developing of new security solutions which will be in great demand with the advent of high-speed next-generation communication systems and networks. This book discusses some of the critical security challenges faced by today's computing world and provides insights to possible mechanisms to defend against these attacks. The book contains sixteen chapters which deal with security and privacy issues in computing and communication networks, quantum cryptography and the evolutionary concepts of cryptography and their applications like chaos-based cryptography and DNA cryptography. It will be useful for researchers, engineers, graduate and doctoral students working in cryptography and security related areas. It will also be useful for faculty members of graduate schools and universities.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821