Authentication of Galileo GNSS Signal by Superimposed Signature with Artificial Noise

Francesco Formaggio, Stefano Tomasin, Gianluca Caparra, Silvia Ceccato, Nicola Laurenti

Department of Information Engineering
University of Padova
Padova, Italy
E-mail: {formaggiof, tomasin, caparrag, ceccatos, nil}@dei.unipd.it

Abstract—Global navigation satellite systems (GNSS) are widely used in many civil applications to get information on position, velocity, and timing (PVT). However, current systems (such as global positioning system (GPS) and Galileo) do not include any feature to authenticate the received signal, therefore leaving open the possibility from an attacker to spoof the GNSS signal and induce a wrong PVT computation at the receiver. In this paper we propose a solution based on the superposition of an authentication message (signature) and artificial noise (AN) on top of the existing navigation signal. Both the authentication message and AN are unpredictable and therefore can not be arbitrarily generated by an attacker. After transmission, through an external public authenticated but asynchronous (thus not useful for PVT) channel, both the authentication message and the AN are revealed, allowing the receiver to check if they were present along the previously received navigation signal. We consider the hypothesis testing problem at the legitimate receiver to decide the authenticity of the message, and we analyze its performance under two attacks: a generation attack in which the attacker does not generate the authentication signal and a replay attack in which a legitimate (including authentication message) signal is replayed by the attacker with a suitable delay in order to induce the desired PVT at the victim. The receiver operating characteristic (ROC) curve is then obtained. As an example, the additional component is described in the context of Galileo E1 Open Service [5], but can trivially be applied to any other GNSS signal.

The paper is organized as follows: in Section II we present the system model while the attack model is described in Section III. In Section IV we derive the analysis for the hypothesis testing of which numerical results are presented in Section V before drawing conclusions in Section VI.

II. System Model

Fig. 1 shows our reference scenario. A satellite (Alice) offers positioning services via a broadcast transmission to both the legitimate receiver (Bob) and the spoofer (Eve). The ground segment communicates with Bob through an authenticated channel, i.e., messages received by Bob over this channel come for sure from the ground segment (rather than from Eve). This channel can be either a terrestrial communication link or a signal component broadcast by the satellite. The information carried by the authenticated channel is available to all users including Eve.

Our model comprises three channels: the authentic navigation channel from the satellite, the authenticated data channel and the attack channel from Eve.

A. Navigation and Attack Channel

The navigation channel connects the satellite to users and is the means through which the legitimate signal $s_L(t)$ transmitted...
We superimpose a synchronous authentication signal where all the users through an authenticated data channel. The code division multiple access (CDMA) using different pseudo-random spread-spectrum sequences called ranging codes. Since the structure of the two components is identical, here we consider only \( E_1 \). The signal carries the unitary power binary data stream \( d_i \) that represents the navigation data for \( E_1 \) with symbol period \( T_s \). The navigation signal can be written as:

\[
    s_N(t) = \sum_{i} d_i s_p(t - iT_s),
\]

where

\[
    s_p(t) = \sum_{i=0}^{N_c-1} c_i u(t - iT_c)
\]

is the spreading pulse with chip period \( T_c = T_s/N_c \), spreading sequence \( c_i = \pm (1/\sqrt{N_c}) \) \( i = 0, \ldots, N_c - 1 \) and unitary-energy chip pulse \( u(t) \). Therefore, \( S_N(t) \) has unitary power. In Galileo OS the chip pulse is a sequence of signed rectangular functions with finite support \( T_c \) [5].

The Galileo Open Service (OS) signal is composed of two signal components: one modulated by the navigation message (the data component \( E_{1B} \)), and one dataless (the pilot component \( E_{1C} \)). The two components are transmitted in code division multiple access (CDMA) using different pseudorandom spread-spectrum sequences called ranging codes. Since the structure of the two components is identical, here we consider only \( E_1 \). The signal carries the unitary power binary data stream \( d_i \) that represents the navigation data for \( E_1 \) with symbol period \( T_s \). The navigation signal can be written as:

\[
    s_N(t) = \sum_{i} d_i s_p(t - iT_s),
\]

where

\[
    s_p(t) = \sum_{i=0}^{N_c-1} c_i u(t - iT_c)
\]

is the spreading pulse with chip period \( T_c = T_s/N_c \), spreading sequence \( c_i = \pm (1/\sqrt{N_c}) \) \( i = 0, \ldots, N_c - 1 \) and unitary-energy chip pulse \( u(t) \). Therefore, \( S_N(t) \) has unitary power. In Galileo OS the chip pulse is a sequence of signed rectangular functions with finite support \( T_c \) [5].

**B. Authenticated Channel**

We assume that the ground segment can communicate with all the users through an authenticated data channel. The authenticated channel is assumed to be of large (infinite) bandwidth and can be for example an Internet connection. We suppose the authentication to be ensured by higher layer authentication protocols. We assume Eve has no control over the information travelling on the authenticated channel and, thus, she can not modify it. This channel is envisioned as a data channel and it is not used directly for ranging purposes.

**C. Authentication Process**

The considered authentication procedure works as follows. We superimpose a synchronous authentication signal \( s_A(t) \) to the ranging signal \( s_N(t) \) at the transmitter so that the transmitted signal is

\[
    s_L(t) = s_N(t) + s_A(t)
\]

\[
    s_A(t) = x(t) + w^*(t),
\]

where \( w^*(t) \) is the artificial noise (AN) component and \( x(t) \) is a the authentication signal that has the same structure of (1) although with a different spreading sequence \( s_{pA}(t) \). \( x(t) \) modulates a binary authentication message of length \( L \) that in vector notation we denote as \( x \). Both \( x(t) \) and \( w^*(t) \) are designed to be orthogonal to \( s_N(t) \) so that the receiver can process navigation signal and authentication component separately. \( x(t) \) and \( w^*(t) \) are also unpredictable, i.e., random generated, however they are revealed to Bob through the authentication channel after the transmission of \( s_L(t) \). Specifically they are revealed at the symbol level, i.e., \( x \) and \( w^* \), where the entries of \( w^* \) are the projections of \( w^* \) into \( s_{pA}(t) \). Bob then subtracts the AN samples and checks the correspondence of the received message with the revealed vector \( x \).

The theoretical analysis carried out in [4] uses concepts of confidential message transmission in physical layer security [6] to design the signaling and analyze its performance. Here instead we follow the hypothesis testing approach, as explained in Section IV.

**D. Channel Model**

In this work we consider a flat one-tap channel with additive white Gaussian noise (AWGN). Under these assumptions the legitimate signal received by Bob is

\[
    r_{B}(t) = g_{AB}(t)s_L(t) + w_B(t),
\]

where \( w_B(t) \) is a Gaussian process with zero mean and spectral density \( N_0(B) \); \( g_{AB}(t) \) is the Alice-Bob channel gain. Similarly, Eve receives

\[
    r_{E}(t) = g_{AE}(t)s_L(t) + w_E(t),
\]

where \( w_E(t) \) is a Gaussian process with zero mean and spectral density \( N_0(E) \) and \( g_{AE}(t) \) is the Alice-Eve channel gain.

**III. ATTACK MODEL**

In the navigation systems, the relevant information extracted by the receiver is the propagation delays estimated by the ranging process. The aim of Eve is to transmit a spoofing signal \( s_E(t) \) that leads Bob to estimate the wrong distance from the satellite.

In absence of the authentication signal component this attack is trivially successful when Eve transmits a replica of \( s_N(t) \), i.e., \( s_E(t) = s_N(t - \tau) \). This is feasible because all the information, such as the spreading code and the navigation message structure needed for generating an arbitrary signal \( s_N(t) \) are publicly available. The introduction of the authentication component \( s_A(t) \) instead would require that Eve generates a spoofing signal that contains also unpredictable information. This has a twofold benefit: a) it forces Eve to observe (at least partially [7], [8]) the transmitted signal in
order to reproduce the unpredictable signal component \( s_A(t) \), and \( b \) because Eve shall observe the signal before reproducing, she can not generate an arbitrary non-causal spoofing signal.

In the following two kinds of attacks are considered:

1) **generation attack**: Eve neglects the authentication component and broadcasts a spoofing signal containing only the navigation component. The signal transmitted by Eve is:

\[
s_E(t) = s_N(t - \tau).
\]

2) **replay attack**: Eve generates a noiseless navigation signal according to the desired range, and a delayed replica of the authentication signal obtained from \( r_E(t) \) given a perfect estimate of \( g_{AE} \). In other words, Eve receives the authenticated message and transmits a suitably delayed replica of it to induce the desired location. The attack signal in this case becomes:

\[
s_E(t) = s_N(t - \tau) + s_A(t - \tau) + w_E(t).
\]

Differently from (6) we note that the replay attack signal is affected by thermal noise \( w_E(t) \). Indeed, as Eve replays a received signal it also contains noise. If this was not the case, the authentication protocol would not work as \( s_E(t) \) would be exactly equal to \( s_L(t) \). Given that, we test the (more realistic) scenario in which Eve’s front end is not ideal.

Several other attacks against navigation signals have been presented in the literature, such as security code estimation and replica (SCER) [9] or forward estimation attacks (FEA) [10]. Given the fact that no information bit of the secret message \( V \) is known a priori, estimation attacks operating at message level, are effective only in obtaining the redundancy of introduced by the channel coding. In this work we consider that the secret message \( x \) is uncoded, therefore FEA attack does not apply. Indeed, coding would be useful also to Eve so we do not see this as a restrictive assumption. The SCER attack does not apply to the authentication signal proposed because, even though the attacker can leverage a matched filter for performing an estimation of \( x(t) \), he cannot do the same on the AN.

Due to the property of auto-correlation of the ranging signal, if the spoofing signal is not aligned within a chip of the spreading code and close in Doppler frequency with the legitimate signal, the two signals will not interfere each other. In this work we assume that the spoofing attack aims at changing significantly the PVT computed by the victim, so we neglect the presence of the legitimate signal when an attack is being performed.

### IV. Detection Statistics

Bob tests the authenticity of the received signal from the demodulated data, i.e., after the tracking loop that performs despreading. In particular, Bob tracks the navigation signal \( s_N(t) \) and buffers the down-converted RF samples and the channel gain estimation \( \hat{g}_B \). When the system reveals the AN samples at a later time, Bob removes them from the stored signal and performs a hypothesis testing on the resulting signal.

In the following analysis bold notation denotes the column vector counterpart of all signals. Moreover, we assume perfect synchronization with the navigation signal (both when it is legitimate and fake). We assume the channel to be constant over \( L \) transmitted symbols. Bob then performs channel estimation every \( L \) samples and his estimate is affected by a zero-mean Gaussian error \( \varepsilon \) with variance \( \sigma^2 \) so that

\[
\hat{g}_B = g + \varepsilon.
\]

#### A. Generation Attack

The aim of Bob is to distinguish through the authentication procedure between the two hypotheses:

\[
\mathcal{H}_0 : r = gx + w_B + (g - \hat{g}_B)w^* \\
\mathcal{H}_1 : r = w_B - \hat{g}_B w^*.
\]

where \( r \) is the vector containing the \( L \) received symbols after the removal of the AN \( w^* \) and \( w_B \) are the i.i.d. noise samples with zero mean and variance \( \sigma^2 \). Hypothesis \( \mathcal{H}_0 \) corresponds to an authentic transmission, while hypothesis \( \mathcal{H}_1 \) corresponds to a transmission without the authentication signal. The authenticity test is performed using binary hypothesis testing by computing the log-likelihood ratio \( \Lambda \) between the probability distribution of \( r \) under both hypotheses:

\[
\Lambda = \frac{p(r; \mathcal{H}_1)}{p(r; \mathcal{H}_0)},
\]

where \( p(\cdot) \) denotes the probability density function. The decision is taken by comparing \( \Lambda \) with threshold \( \gamma \):

\[
\begin{cases} 
\text{decide } \mathcal{H}_1 \text{ if } \Lambda > \gamma \\
\text{decide } \mathcal{H}_0 \text{ if } \Lambda < \gamma.
\end{cases}
\]

We can now define the probability of false alarm and miss detection as

\[
P_{FA} = P[\Lambda > \gamma; \mathcal{H}_0] \\
P_{MD} = P[\Lambda < \gamma; \mathcal{H}_1],
\]

where \( P[\cdot] \) is the probability function.

In general, the distribution of \( \Lambda \) is not known in closed form. However, \( \Lambda \) is the ratio between two Gaussian pdfs. This is because \( x \) and \( w^* \) are revealed to Bob and therefore are to be considered constant values. Then, under \( \mathcal{H}_0 \) \( r \) is a Gaussian vector with pdf:

\[
p(r; \mathcal{H}_0) \sim \mathcal{N}(\hat{g}_B x, X\sigma^2 + \sigma^2 w_B I),
\]

where

\[
X = (x + w^*)(x + w^*)^T.
\]

Similarly, under \( \mathcal{H}_1 \)

\[
p(r; \mathcal{H}_1) \sim \mathcal{N}(-\hat{g}_B w^*, \sigma^2 w_B I).
\]

If there is no estimation error, i.e., \( \varepsilon = \sigma^2 = 0 \), it is possible to derive a closed-form expression for the distribution of the log-likelihood ratio. Taking the logarithm of \( \Lambda \) and doing simple algebraic calculations, the two \( ||r||^2 \) terms cancel out leaving with a linear combination of the scalar entries in \( r \). Since in both \( \mathcal{H}_0 \) and \( \mathcal{H}_1, r \) is an uncorrelated Gaussian vector, also
The log-likelihood ratio \( \Lambda \) is normally distributed, and for the two hypotheses we have:

\[
\begin{align*}
\mathcal{H}_0 & : \log \Lambda \sim \mathcal{N}(\mu_A, \sigma^2_A) \\
\mathcal{H}_1 & : \log \Lambda \sim \mathcal{N}(\mu_A, \sigma^2_A),
\end{align*}
\]

where

\[
\mu_A = \frac{g_B^2}{2\sigma_w^2}||x + w^*||^2, \quad \sigma^2_A = \frac{g_B^2}{2\sigma_w^2}||x + w^*||^2.
\]

We can compute \( P_{MD} \) given a fixed \( P_{FA} \):

\[
\gamma' = Q^{-1}(P_{FA}), \quad P_{MD} = 1 - Q(\gamma''),
\]

where \( \gamma' \) and \( \gamma'' \) are the results of the normalizations needed to use the standard Gaussian tail distribution function \( Q(\cdot) \).

Note that from (17) the balancing between the power of \( x \) and \( w^* \) is not relevant, as only the norm square of the sum matters. However, the unpredictability of the AN is still needed to avoid Security Code Estimation and Replay (SCER) attacks, while the binary message \( x \) does not suffer from the quantization issues discussed in [4].

B. Replay Attack

In case of replay attack the two hypotheses are:

\[
\begin{align*}
\mathcal{H}_0 & : \gamma = gx + \gamma_B + (g - \gamma_B)w^* \\
\mathcal{H}_1 & : \gamma = g(x + \gamma_B) + \gamma_B + (g - \gamma_B)w^*,
\end{align*}
\]

where \( \gamma_B \) is the zero-mean Gaussian noise introduced by Eve’s front end.

The distribution of \( \gamma \) is known in closed form only under \( \mathcal{H}_0 \) and it is given by (13), while \( \mathcal{H}_1 \) requires numerical integration by conditioning over \( g \) and applying the total probability theorem:

\[
p(\gamma; \mathcal{H}_1) = \int_{-\infty}^{+\infty} p(\gamma|g; \mathcal{H}_1)p(g)dg,
\]

where from (8)

\[
g \sim \mathcal{N}(\gamma_B, \sigma^2).
\]

Therefore in this case \( p_{FA} \) and \( p_{MD} \) can not be obtained in closed form.

For the special case \( \varepsilon = \sigma^2 = 0 \), the distribution of \( \log \Lambda \) is available in closed form also in this case. Since \( g = \gamma_B \), \( \gamma \) becomes again a Gaussian vector, the two hypotheses become

\[
\begin{align*}
\mathcal{H}_0 & : \gamma \sim \mathcal{N}(\gamma_B, \sigma^2) \\
\mathcal{H}_1 & : \gamma \sim \mathcal{N}(\gamma_B, \gamma_B^2\sigma^2 + \sigma^2).
\end{align*}
\]

If we define \( \gamma' = \gamma - \gamma_Bx \) we can apply the results in [11] and write the test statistic

\[
\Lambda' = \frac{\sigma^2 + \gamma_B^2\sigma^2}{2(\gamma_B^2\sigma^2 + \gamma_B^2\sigma^2)}||\gamma'||^2,
\]

which can be derived by taking the log of (10), scaling and ignoring the non-\( \gamma' \)-dependent terms. The entries of \( \gamma' \) are independent Gaussian random variables and therefore

\[
\Lambda' \sim \chi^2_L.
\]

The same procedure of (18) can be carried out using the cumulative distribution function of the Chi-square random variable instead of \( Q(\cdot) \).

V. Numerical Results

A. Generation Attack

In Fig. 2 we show the ROC of the hypothesis testing problem under generation attack for the case \( \varepsilon = 0 \) following the procedure in (18). We set \( \sigma_w^2 = -11 \) dB, as it corresponds to 35 dB of carrier-to-noise ratio \( (C/N_0) \), a typical value for navigation systems. Moreover we set \( L = 5 \), which for the Galileo system corresponds to 20 ms of signal observation.

Note that the defense against this type of attack is easy as Eve does not transmit any authentication component. Indeed, even with \( L = 5 \) we can obtain a miss detection probability of \( 10^{-10} \) for a false alarm probability of \( 4 \times 10^{-7} \).

Eve can still induce a lower channel estimate to Bob and as a result he will not be able to distinguish between \( \mathcal{H}_0 \) and \( \mathcal{H}_1 \), since the distributions of the two received signals become closer as \( g \) decreases. This results in worse \( p_{FA} \) and \( p_{MD} \) curves, as shown in Fig. 2. However, \( g \) can not be too low, otherwise Bob would not be able to track the signal, leaving Eve with an unsuccessful attack.

The same considerations hold for Fig. 3 showing the ROC for a scenario taking into account also the estimation noise variance \( \sigma^2 \). In this case the noisy estimate \( \gamma_B \) impacts on both the false alarm and miss detection performance.

B. Replay Attack

The replay attack is a powerful attack since Eve sends an exact copy (plus the noise \( w_E(t) \)) of the received signal, together with the authentication component. The two hypotheses are then more difficult to distinguish because, we recall, they differ only in the thermal noise \( w_E \).

In Fig. 4 we show the ROC using the procedure described in Section IV-B. For this scenario what matters is not the absolute value of \( \sigma^2 \), but its ratio with \( \sigma^2 \). \( P_{FA} \) and \( P_{MD} \) are then shown as a function of the \( C/N_0 \) gap between Bob and Eve, where for a gap larger than 0 dB Eve has a physical advantage over Bob, i.e., \( \sigma^2 < \sigma^2 \). We see that in order to obtain an
acceptable performance, either Eve has the same channel as Bob, or we need a larger $L$, as we show next.

Fig. 5 shows the ROC curve obtained evaluating numerically the integral in (20) to take into account also in this case the estimation noise power, which has been set to $\sigma^2_{\varepsilon} = 10^{-11} \sigma^2_{w_B}$. We see how it is always possible to improve performance by increasing $L$ which, however, is now at least 150 symbol long, corresponding to 600 ms. Note that defending against the generation attack is easier as it requires a shorter observation time.

VI. CONCLUSIONS

In this paper we provided an additional analytical analysis of the authentication protocol proposed in [4]. The analysis has been carried out using binary hypothesis testing and considering two types of spoofing attacks: a simple generation attack and a replay attack.

With respect of the original formulation of the protocol we introduced the presence of the channel gain and took into account the channel estimation error. We showed how using the authentication protocol against the generation attack requires a small amount of authentication symbols, while the replay attack is more powerful and requires the transmission of a longer authentication message.

REFERENCES