Use and Validation of Supervised Machine Learning Approach for Detection of GNSS Signal Spoofing

Silvio Semanjski
Department of Communication, Information, Systems & Sensors
Royal Military Academy
Brussels, Belgium
silvio.semanjski@rma.ac.be

Alain Muls
Department of Communication, Information, Systems & Sensors
Royal Military Academy
Brussels, Belgium
alain.muls@rma.ac.be

Ivana Semanjski[1,2]
[1]*Department of Industrial Systems Engineering and Product Design*
Ghent University
Ghent, Belgium
[2]*Industrial Systems Engineering (ISyE), Flanders Make, www.FlandersMake.be*
Ghent University,
Ghent, Belgium
ivana.semanjski@ugent.be

Wim De Wilde
Septentrio N.V.
Septentrio N.V.
Leuven, Belgium
wim.dewilde@septentrio.com

***Abstract*—Spoofing of the GNSS signals presents continuous threat to the users of safety of life applications due to unaware use of false signals in generating position and timing solution. Among numerous anti-spoofing techniques applied at different stages of the signal processing, we present approach of monitoring the cross-correlation of multiple GNSS observables and measurements as an input for supervised machine learning based approach to detect potentially spoofed GNSS signals. Both synthetic, generated in laboratory, and real-world spoofing datasets were used for verification and validation of the supervised machine learning algorithms for detection of the GNSS spoofing.**

Keywords—Global Navigation Satellite System, Spoofing, Support Vector Machines, Principal component analysis, Safety-of-Life, Position-Navigation-Timing, GPS, GNSS, PNT, SVM, SOL

I. Introduction

The unencrypted, open services of the GNSS core constellations (GPS, GALILEO, GLONASS, BEIDOU) being used in Safety-of-Life (SoL) applications such as aircraft navigation or time synchronization of traffic control systems, are particularly vulnerable to the signal spoofing due to risk of unaware use of the manipulated data for Position-Navigation-Timing (PNT), which is of continuous interest to the industry [1]. The threat of GNSS signal spoofing proliferates with advances of the digital signal processing and transmitting capable Software Defined Radio (SDR) type of spoofers, which poses significant challenge to developers of spoofing detection.

The GNSS measurements data produced at the stage of generating Position-Velocity-Time (PVT) solution within receiver, contains number of observables which can be monitored and measurements cross-correlated with purpose of spoofing detection.

One of the spoofing methods exploits GNSS receiver clock offset by manipulating the clock drift (time derivative of the clock offset) estimates, resulting in deviations of the pseudorange measurements and consequently affecting accuracy of the PVT solution. However, associated power level adjustment of the spoofing signal is required to compete with genuine signals when acquired and tracked by the GNSS receiver, resulting in slightly elevated C/N_0. Monitoring of the C/N_0 for detection of the spoofed signals overtaking authentic ones proves as a weak technique due to function of the Automatic Gain Control (AGC) which compensates for fluctuation of the power level of the received signal, knowing that only 1.1 dB of power ratio between spoofed and authentic signal is sufficient to force the receiver to lock onto the spoofed signal [1].

Contrary to monitoring of single (or few) estimates within receiver, the approach to monitor cross-correlation among all available and statistically significant GNSS observables and measurements is proposed as a detection of the spoofing signal presence.

In this paper we present results of the verification and validation of machine learning based detection of the GNSS signal spoofing, in particular applying the Support Vector Machine (SVM) classification. The imposed conditions for the verification were success rate and repeatability of the classification among three synthetically generated spoofing datasets, while the successful classification using real-world spoofing event dataset was chosen for the successful outcome for the validation of the SVM-based spoofing detection.

II. Data and Methods

A. Data collection

Different spoofing datasets have been used in development and validation of the SVM-based GNSS signal spoofing detection; synthetically generated (simulated) ones, and those resulted from real-world spoofing event. In our work we have used the measurements of the L1 and L2 C/A code signals at the output rate of 1 Hz contained in 'Measurement Epoch' and 'Measurement Extra', and 'PVT Geodetic' measurements blocks for both synthetically generated and real-world spoofing datasets. The level of dynamics (the balance between noise and dynamics in the GNSS measurements and the PVT solution) the GNSS antenna is subjected to during aforementioned measurements was set to high (in which case, high-frequency motion becomes visible at the expense of an increase in the noise).

Those three spoofing datasets used for machine learning algorithms development and verification, have been generated with matching power levels by modified Spirent GNSS signal and constellation simulator connected with the transmitting antenna placed in Wave Field Synthesis (WFS) anechoic chamber at the Fraunhofer FORTE facility [5][6]. In each dataset six channels are representing "authentic" (non-manipulated) GPS C/A code signal at L1 frequency, while the other six channels are representing spoofed version of the

978-1-7281-2445-2/19/$31.00 ©2019 IEEE

signal. In all three datasets spoofing gets enabled 120 seconds into the test scenario, which allows for both spoofed and non-spoofed epochs to be present in each dataset. In our spoofing scenarios, which have been previously used [3], an intermediate timing attack consists of hijacking of the Pulse-Per-Second (PPS) output of the receiver through programmed clock divergence with increase in Carrier-to-Noise Density ratio (C/N_0) of 2 dB or more for each tracked satellite. Three subsets of programmed clock divergence have been created: with 5 ns/s, 1 ns/s and 0.3 ns/s rate of time pulling. Three spoofing datasets used were results of measurements (with 948, 1493, and 4665 total epochs, respectively) made by Septentrio AsteRx-U GNSS receiver capable of tracking all known signals of multiple constellations, which logged extensive number of GNSS observables including baseband samples in Septentrio Binary Format (SBF) blocks.

The actual, the real-world spoofing event caused by unintentional re-radiation of the GNSS signal over-the-air was used as a dataset for validating the GNSS spoofing detection method. The record of this event was captured by the nearby GNSS receiver and we use it as an independently created validation dataset (was not a part of the model training). This dataset contains the following mixed authentic and spoofed signals: GPS L1 C/A, GPS L2 P, GPS L2C, QZSS L1 C/A, QZSS L2 C, GLONASS L1 C/A, GLONASS L2 C/A, IRNSS L5, GALILEO E1, GALILEO E1a, GALILEO E1b, MSS L band, SBAS L1 C/A, BeiDou B1I, and BeiDou B2. Further on we will refer to this dataset as the unintentional spoofing dataset and it was produced by measurements (with 459 total epochs) in the field by Septentrio AsteRx-m2 GNSS receiver.

B. Corelation analysis

As a part of the data pre-processing step, we examined the correlation among all the available variables and the indication if the GNSS signal was spoofed or not. For correlation analysis, we used the well-known Pearson correlation [7], [8]:

$$r_{x,y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (1)$$

where:
- n is sample size
- x_i, y_i are the individual sample points indexed with i
- $\bar{x}$ and $\bar{y}$ are the sample means

As the input for the next step (C-SVM) we only considered variables that have statistically significant correlation with the indication if the GNSS signal was spoofed or not.

C. Support Vector Machines classification

To detect GNSS signal spoofing we selected the C-Support Vector Machines (C-SVM) based approach. Reason for this is the evidence of the SVM's high applicability to diverse domains, both for classification [9], [10] or regression analysis [11],[12]. Furthermore, there is already provisionally supported hypothesis that C-SVM has high potential in assisting the detection of GNSS signal spoofing attempts [13]. Additionally, we wanted to obtain scalable runtime in regard to the number of input samples, for which literature suggests that C-SVM is a better option over, for example, nu-SVM classification [4].

To prepare our data set, we firstly divided our dataset in two parts: the training (Z_1) and the test dataset (Z_2). This division is made based on the 70:30 principle, randomly sorting the 70% of data into the training set and 30% into the test set. To map the multiclass problem into binary classification problem, we use one-against-all approach and minimization error function [10]:

$$\frac{1}{2}w^T w + C\sum_{i=1}^{N}\xi_i \quad (2)$$

which is a subject to the constraints:

$$y_i(w^T\,\phi(x_i) + b) \geq 1 - \xi_i \quad (3)$$

$$\xi_i \geq 0 \quad (4)$$

Where:
$i = 1, \ldots, N$,
w - the vector of coefficients;
C - the capacity constant;
b – constant;
ξ_i - parameters for handling non-separable data (inputs).

The index i labels the N training cases. y_i ($y \in \pm 1$) represents the class labels and x_i represents the model's independent variables. The ϕ indicates the kernel function that transforms input (independent variables) to the feature space. In our example we used the radial basis function (RBF) where γ is an adjustable parameter of the kernel function:

$$K(X_i, X_j) = \phi(X_i)\cdot\phi(X_j) = \exp\left(-\gamma\left|X_i - X_j\right|^2\right) \quad (5)$$

The reader can find more details on the RBF in literature [14], [15], [16]. The values of capacity constants C (2) and γ (5) are important to keep the training error small and in order to generalize well [17]. However, since it is not possible to known upfront the best values of these constrains, we applied the incremental grid-search. For C, in range from 1 to 20, with the step equal to 1, and for γ, in range from 0 to 0.5, with the step equal to 0.01. The values with the best average 10-fold cross-validation accuracy were chosen to be further used on the test data. The obtained values were 19 for C and 0.71 for γ.

For the v-fold cross-validation, the total number of cases was randomly divided into v sub-samples, where $v = 10$, $Z_1, Z_2, \ldots, Z_v$ of equal sizes ($N_1, N_2, \ldots, N_v$). The v-fold cross-validation estimate (6) is the proportion of cases in the subsample Z that are being misclassified by the classifier constructed from the subsample $Z - Z_v$:

$$R\left(d^{(v)}\right) = \frac{1}{N_v}\sum_{(x_n, j_n)\in Z_v} X(d^{(v)}(x_n) \neq j_n) \quad (6)$$

Where:
$-\ d^{(v)}(x)$ is the classifier calculated from the sub sample $Z - Z_v$
$-$ X is the indicator function for which the following is valid:

$X = 1$, if the statement $X\left(d^{(v)}\right) \neq j_n)$ is true (7)

$X = 0$, if the statement $X\left(d^{(v)}\right) \neq j_n)$ is false. (8)

Following this, we consider the test sample estimate to be the proportion of cases in the test dataset that are misclassified by the classifier constructed from the learning dataset. This estimate is computed in the following way:

$$R(d) = \frac{1}{N_2}\sum_{(x_n, j_n)\in Z_2} X(d(x_n) \neq j_n) \quad (9)$$

In the next step, we considered completely independent dataset, which was not use nor for training nor for testing, where unintended spoofing occurred in uncontrolled environment. We applied parameters obtained from the training step in order to validate the C-SVM constructed C-SVM.

D. Principal component analysis

To better understand the relations between the selected variables and the indication if the GNSS signal was spoofed or not, we implemented the factor analysis, or more precisely the Principal Component Analysis (PCA).

The PCA is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components [18]. This transformation is conducted in such a way that the first principal component captures the largest possible variance, and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding ones.

III. Results

A. Corelation analysis

The table below shows the partial correlations among the following variables:

(1) Lock time [s]
(2) C/N_0 [0.25 dB-Hz]
(3) Pseudorange [m]
(4) Carrier Doppler frequency [0.0001 Hz]
(5) Full carrier phase [cycles]
(6) Multipath correction [0.001 m]
(7) Code variance [0.0001 m^2]
(8) Carrier variance [$mcycle^2$]
(9) Carrier multipath correction [1/512 cycle]
(10) Receiver clock bias [ms]
(11) Receiver clock drift [ppm]
(12) Spoofing indication

The blue color in TABLE I. indicates the values of the statistically significant ($p<0.05$) correlation.

TABLE I. Correlation Matrix

	1	2	3	4	5	6	7	8	9	10	11	12
1	1.00	-0.02	0.28	-0.27	0.28	-0.03	-0.10	-0.29	-0.03	0.08	-0.18	-0.33
2	-0.02	1.00	-0.34	0.22	-0.34	-0.38	-0.66	-0.36	-0.35	-0.02	0.03	0.07
3	0.28	-0.34	1.00	-0.44	1.00	0.00	0.19	0.07	-0.01	0.11	-0.05	-0.10
4	-0.27	0.22	-0.44	1.00	-0.44	0.01	-0.09	0.00	0.01	-0.02	0.05	0.09
5	0.28	-0.34	1.00	-0.44	1.00	0.00	0.19	0.07	-0.01	0.11	-0.05	-0.10
6	-0.03	-0.38	0.00	0.01	0.00	1.00	0.14	0.09	0.11	0.01	0.02	0.01
7	-0.10	-0.66	0.19	-0.09	0.19	0.14	1.00	0.64	0.59	-0.02	0.00	0.05
8	-0.29	-0.36	0.07	0.00	0.07	0.09	0.64	1.00	0.34	-0.06	0.24	0.59
9	-0.03	-0.35	-0.01	0.01	-0.01	0.11	0.59	0.34	1.00	-0.01	0.01	0.02
10	0.08	-0.02	0.11	-0.02	0.11	0.01	-0.02	-0.06	-0.01	1.00	-0.07	-0.20
11	-0.18	0.03	-0.05	0.05	-0.05	0.02	0.00	0.24	0.01	-0.07	1.00	0.47
12	-0.33	0.07	-0.10	0.09	-0.10	0.01	0.05	0.59	0.02	-0.20	0.47	1.00

B. Data exploratory analysis

The training of the model has been referenced to the receiver clock drift and C/N_0, reflected from three different programmed clock divergence values, with 5 ns/s, 1 ns/s and 0.3 ns/s rate of time pulling, respectively. In Figure 1, Figure 3, and Figure 2, the receiver clock drift is shown for each of aforementioned clock divergence values.

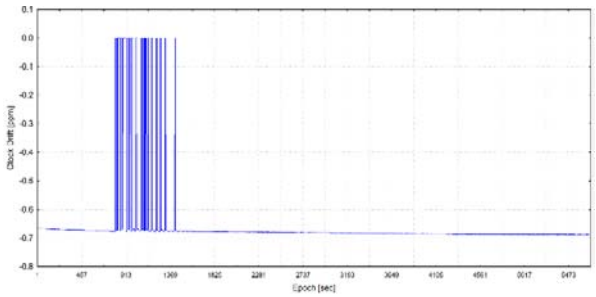

Figure 1 Receiver clock drift from training dataset with programmed 5 ns/s of time pulling

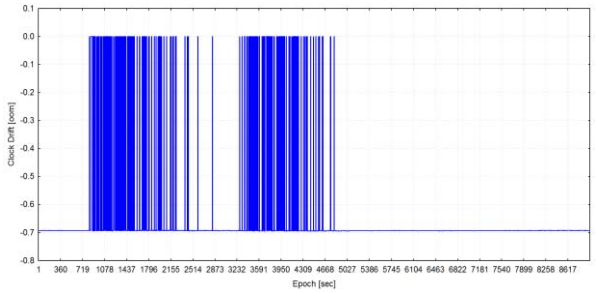

Figure 2 Receiver clock drift from training dataset with programmed 1 ns/s of time pulling

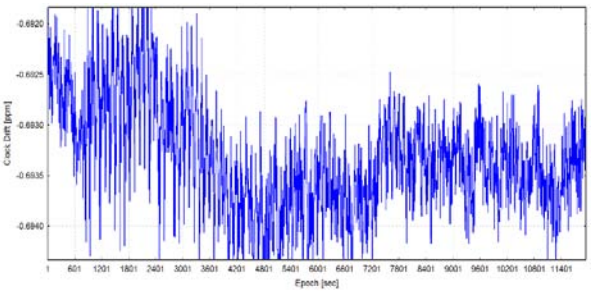

Figure 3 Receiver clock drift from training dataset with programmed 0.3 ns/s of time pulling

In Figure 4, the receiver clock drift for real-world spoofing event is shown.

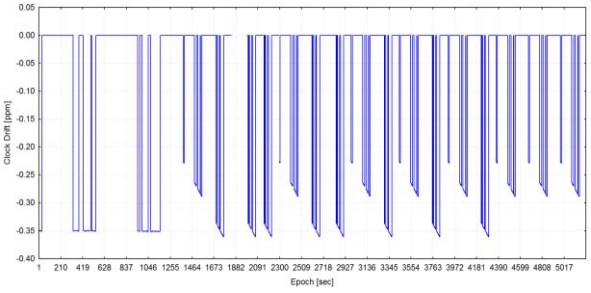

Figure 4 Receiver clock drift from real-world spoofing dataset (validation dataset)

In Figure 5, the Carrier-to-Noise density ratio of satellites used (SVs: G01, G03, G06, G19, G20, G23), is shown superimposed for all three training datasets. Start of the

spoofing event is visible on the graph as jump in dB-Hz value 120 seconds into the test.

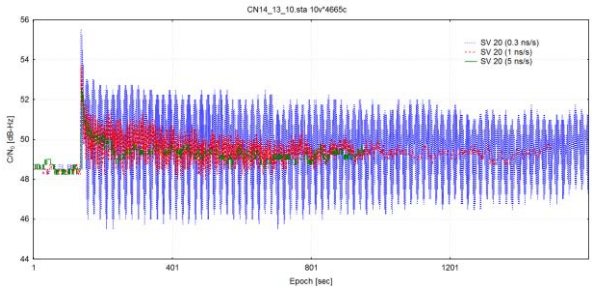

Figure 5 C/N$_0$ of SV 20 from all three training datasets

In Figure 6, the Carrier-to-Noise density ratio for SV 20 from real-world spoofing dataset is shown.

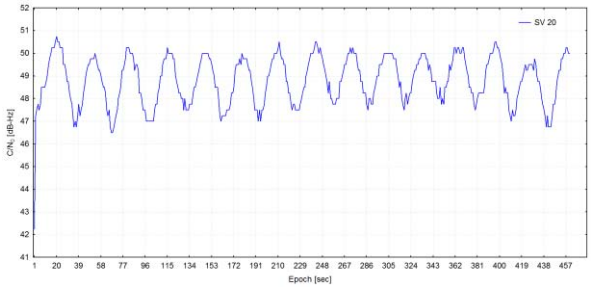

Figure 6 C/N$_0$ of SV 20 from real spoofing dataset

C. *Support vector machines*

Table II gives a summary of the results for the applied C-SVM method. The overall success rate of the proposed approach was 97.8%, whereas the cross-validation error was slightly higher (97.8%). In total, 924 support vectors were used, of which 456 belonged among the "authentic" GNSS observations and reminder among the spoofed GNSS signal observations. Also, when the trained model was applied to the independent, unintentional spoofing, dataset the achieved success rate was even higher, indicating complete detection of the unintentional spoofing in uncontrolled environment.

Considering the confusion matrix shown in Table III, 1.5% of the authentic GNSS signals were confused to be spoofed and 7.6% of spoofed signals were confused to be authentic in the test dataset. However, for the independent dataset (Table IV) none of the confusions appeared.

TABLE II. MODEL SUMMARRY

	Value
Number of independents	11
SVM type	Classification type 1
Kernel type	Radial Basis Function
Number of SVs	57 (52 bounded)
Number of SVs (authentic GNSS signal)	465
Number of SVs (spoofed GNSS signal)	459
Cross -validation accuracy	97.8 %
Class accuracy (training dataset)	97.7 %
Class accuracy (test dataset)	97.8 %
Class accuracy (overall)	97.8 %
Class accuracy (independent dataset)	100.0%

TABLE III. CONFUSION MATRIX FOR TEST DATASET

	Authentic GNSS signal	Spoofed GNSS signal
Authentic GNSS signal	5587	88
Spoofed GNSS signal	55	663

TABLE IV. CONFUSION MATRIX FOR INDEPENDENT DATASET

	Authentic GNSS signal	Spoofed GNSS signal
Authentic GNSS signal	363	0
Spoofed GNSS signal	0	4686

D. *Principal component analysis*

The PCA indicated 10 principal components. Figure 7 shows these components and the successive eigenvalues on the *x*-axis. One can notice that the first component accounts for almost 37% of the spoofing signal indication variations, the second for 21% and the third for 11%. Finally, the last component captures for 1.7% of the variations. Having this in mind, we wanted to examine which variables contributed to which factors. Table V shows the relation between the components (columns) and variables used to build the model (rows). One can see that the first component is built mainly based on the C/N$_0$, pseudorange, full carrier phase, and code variance variables (the respective eigenvalue is most correlated with these variables). Also, all of the above mentioned variables have high negative correlations, with the exception of C/N$_0$.

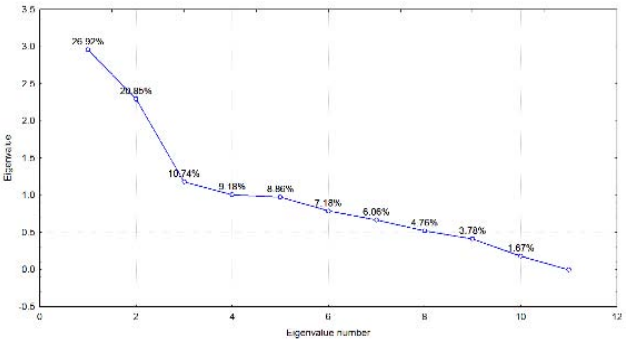

Figure 7 Principal components

TABLE V. FACTOR COORDINATES OF VARIABLES

Factors / Variables	1	2	3	4	5	6	7	8	9	10
1	-0.17	0.57	-0.34	0.10	-0.19	0.55	-0.20	0.35	0.06	0.01
2	0.77	0.25	0.23	0.21	0.03	0.06	-0.03	-0.11	0.44	-0.19
3	-0.73	0.56	0.22	-0.03	0.06	-0.16	-0.22	-0.09	0.06	-0.01
4	0.48	-0.44	-0.10	0.03	0.21	-0.19	-0.69	0.12	-0.05	0.01
5	-0.73	0.56	0.22	-0.03	0.06	-0.16	-0.22	-0.09	0.06	-0.01
6	-0.24	-0.24	-0.42	-0.80	0.06	0.03	-0.02	-0.06	0.24	-0.06
7	-0.73	-0.53	-0.07	0.22	-0.04	-0.02	0.02	0.12	-0.09	-0.33
8	-0.49	-0.60	0.32	0.11	0.08	-0.07	0.09	0.36	0.33	0.15
9	-0.44	-0.51	-0.25	0.35	-0.09	0.31	-0.14	-0.45	0.13	0.10
10	-0.08	0.19	-0.26	0.17	0.91	0.11	0.15	0.01	-0.01	0.00
11	0.01	-0.26	0.70	-0.31	0.19	0.53	-0.10	-0.06	-0.14	-0.03

Figure 8 displays the component coordinates for the first two principal components in the unit circle. Because the analysis is based on correlations, the largest component coordinate (variable-component correlation) that can occur is equal to 1. Furthermore, the sum of all squared principal component coordinates for a variable (i.e. squared correlations between the variable and all components) cannot exceed one. Hence, all component coordinates must fall within the unit circle indicated in the graph. Following this analogy, the circle provides a visual indication (scale) of how well each variable is represented by the current set of principal components. The closer a variable in this plot is located to the unit circle, the better is its representation by the current coordinate system.

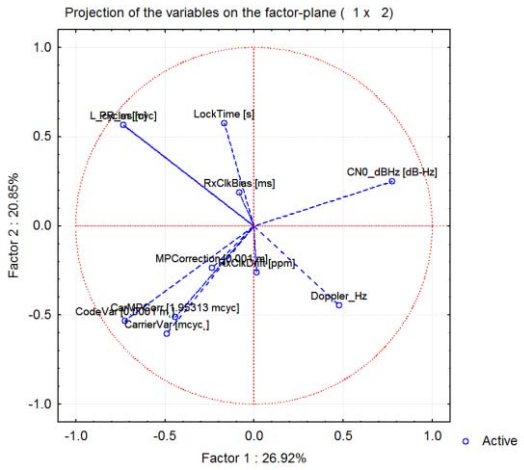

Figure 8 First two factors plot

IV. Discussion

Since the training and test of the SVM was performed on all three synthetic spoofing datasets joined together, the impact of the most intentional clock drift manipulation values on the classification was examined. In particular, the one of three datasets with time-pulling attack of 0.3 ns/s was chosen for this test. In this dataset the clock drift of the spoofer with respect to the authentic signal was only 0.3 ns/s, which is 0.3 ppb (0.0003 ppm) what is smaller than the noise on the clock drift and close to the Allan deviation of the oscillator. In this case spoofing detection relying on observing the clock drift as the independent variable would not be possible (although observable in its C/N_0, which much too regular for a "real" signal, because there's a systematic succession of constructive and destructive interference between the spoofer and the authentic signal). The amplitude decreases over time as the spoofing chip moves apart from the authentic chip.

However, correlation results shows that even at this level of the clock drift manipulation, is possible to perform classification of the spoofed and authentic signal variables.

The correlation matrix indicates the statistically relevant correlation among the variables selected for our model (significant at $p < 0.05$) and that the selected variables can give a good indication whether the observed signal is spoofed or not. Furthermore, principal component analysis resulted in ten principal components out of 11 variables, indicating the high value of the selected variables as the dimensionality of the model would not significantly change with using the components over the variables included in the model. When having a look at the principal component's eigenvalues, it is indicative that for the first three components the percentage of the spoofing signal indication variations falls rapidly and this slope becomes less steep afterwards. Hence, one could argue that the first three components capture the most of the variability and that the contribution of each following one is less and less relevant.

By having a look at the SVM results, one can notice quite big number of the supporting vectors. This indicates large complexity of the existing model and potential overfitting. However, having in mind that the approach was validated on an independent (unintentionally spoofed) dataset and achieved high success rate, the overfitting seems not to be the case. Nonetheless, the high complexity still remains a challenge that needs to be tackled in future research as it can result in long computing time, hence hinder the applicability of the approach for the real time sensitive applications among which any Safety-of-Life navigation application can be accounted for.

V. Conclusion

The obtained results lead to several conclusions. For one, correlation analysis seems as good approach for selection of variables that serve as an input for the supervised machine learning approach. In our example, the principal component analysis did not result in major reduction of the model dimensionality when applied. However, the PCA analysis seems as a beneficial complementary analysis to better understand relations among the selected variables. Furthermore, the SVM seem to be promising approach to detect potential GNSS signal spoofing attempts. In our example, the SVM resulted in a quite extensive number of the supporting vectors, indicating that the separation among spoofed and authentic signals is a quite complex task. However, the validation results on an unintentionally spoofed set of observations seem to speak on behalf of the proposed approach as a robust one for future applications.

References

[1] D. Schmidt, K. Radke, S. Camtepe, E. Foo, and M. Ren, "A Survey and Analysis of the GNSS Spoofing Threat and Countermeasures," ACM Comput. Surv., 2016.

[2] N. O. Tippenhauer, C. Pöpper, K. B. Rasmussen, and S. Capkun, "On the requirements for successful GPS spoofing attacks," in Proceedings of the 18th ACM conference on Computer and communications security - CCS '11, 2011.

[3] G. S. De Wilde, Wim; Sleewaegen, Jean-Marie; Bougard, Bruno; Cuypers, Gert; Popugaev, Alexander; Landmann, Markus; Schirmer, Christopher; Roca, Daniel Egea; López-Salcedo, José A.; Granados, "Authentication by Polarization: A Powerful Anti-Spoofing Method," in 31st International Proceedings of the 31st International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2018), 2018, pp. 3643–3658.

[4] C.-C. Chang and C.-J. Lin, "Training v -Support Vector Classifiers: Theory and Algorithms," Neural Comput., vol. 13, no. 9, pp. 2119–2147, 2001.

[5] M. Rügamer, A.; Del Galdo, G.; Mahr, J.; Rohmer, G.; Siegert, G.; Landmann, "Over-The-Air Testing using Wave-Field Synthesis," in Proceedings of the 26th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2013), 2013, pp. 1931–1943.

[6] C. Schirmer, M. H. Landmann, W. A. T. Kotterman, M. Hein, R. S. Thoma, G. Del Galdo, and A. Heuberger, "3D wave-field synthesis for testing of radio devices," in 8th European Conference on Antennas and Propagation, EuCAP 2014, 2014, pp. 3394–3398.

[7] J. Wang, "Pearson Correlation Coefficient," Encycl. Syst. Biol., no. 1, pp. 1671–1671, 2013.

[8] J. Adler and I. Parmryd, “Quantifying colocalization by correlation: The pearson correlation coefficient is superior to the Mander’s overlap coefficient,” Cytom. Part A, vol. 77, no. 8, pp. 733–742, 2010.

[9] S. Joo, C. Oh, E. Jeong, and G. Lee, “Categorizing bicycling environments using GPS-based public bicycle speed data,” Transp. Res. Part C Emerg. Technol., vol. 56, pp. 239–250, Jul. 2015.

[10] I. Semanjski and S. Gautama, “Crowdsourcing mobility insights – Reflection of attitude based segments on high resolution mobility behaviour data,” Transp. Res. Part C Emerg. Technol., vol. 71, 2016.

[11] E. I. Vlahogianni, “Optimization of traffic forecasting: Intelligent surrogate modeling,” Transp. Res. Part C Emerg. Technol., vol. 55, pp. 14–23, Jun. 2015.

[12] J. Wang and Q. Shi, “Short-term traffic speed forecasting hybrid model based on Chaos – Wavelet Analysis-Support Vector Machine theory,” Transp. Res. Part C, vol. 27, pp. 219–232, 2013.

[13] S. Semanjski, I. Semanjski, W. De Wilde, and A. Muls, “Cyber-threats Analytics for Detection of GNSS Spoofing,” in Data Analytics 2018, 2018, no. c, pp. 136–140.

[14] D. S. Broomhead, D. Lowe, "Radial basis functions, multi-variable functional interpolation and adaptive networks", Royal signals and radar establishment malvern, Technical report, 1988.

[15] B. Schölkopf, K-K. Sung, C. Burges, F. Girosi, P. Niyogi, T. Poggio, V. Vladimir, "Comparing Support Vector Machines with Gaussian Kernels to Radial Basis Function Classifiers", Transactions on Signal Processing, vol. 54, no.11, pp. 2758–2765, 1997.

[16] O. Chapelle, V. Vapnik, "Model Selection for Support Vector Machines", in NIPS'99 Proceedings of the 12th International Conference on Neural Information Processing Systems, 1999, p.p. 230-236.

[17] D. Anguita and L. Oneto, “In-sample Model Selection for Support Vector Machines,” in The 2011 International Joint Conference on Neural Networks, 2011, p. 2011.

[18] S. Wold, K. Esbensen, and P. Geladi, “Principal component analysis,” Chemom. Intell. Lab. Syst., vol. 2, no. June 2001, pp. 37–52, 1987.