Emitting Devices Identification Based on Variational Mode Decomposition of Captured Bluetooth Signals

Alghannai Aghnaiya[1]
[1] Department of Communications,
College of Electronic Technology,
Bani Walid, Libya
[1] elganai.1962@gmail.com

Saleh Abofana[2]
[2] Department of Computer,
Faculty of Science, University
Bani Walid, Libya
[2] s.g.abofana@gmail.com

Ismael Salem[3]
[3] Department of Communications,
College of Electronic Technology,
Bani Walid, Libya
[3] ismael.salem@yahoo.com

***Abstract*—Recently, the use of communication systems is increased therefore the security networks has become more important. Radio frequency fingerprinting (RFF) is one of the communication networks security techniques based on identification of the unique features of the RF transient signals. Transient signal detection is a very important part in the classification system. If a transient signal is not detected precisely, then the features extracted from the received signal will not show the correct unique features of the transmitting devices. For this purpose, variational mode decomposition is applied to the recorded Bluetooth (BT) signal before the transient detection process. In this paper, VMD is applied before transient detection as denoying method in RF fingerprinting of Bluetooth device identification to increase the accuracy of transient detection and the classification rate. The method has three stages: Firstly, VMD is used to decompose the recorded BT signal into a series of band-limited modes. Then, the selected modes are added to reconstruct the original BT signal. Secondly, transient detection and different statistical features are extracted from the transients such as variance, kurtosis and skewness. Finally, a linear support vector machine (LSVM) classifier is used to identify 20 mobile phones under different SNR ranges. 20 mobile phones with 150 transients each were tested in this work. Results show that the proposed VMD identification method based on the decomposition of recorded signal outperforms the existing VMD method based on decomposition of transient signal for different ranges of SNR.**

Keywords—Variational mode decomposition, Bluetooth signals, specific emitter identification, feature extraction, signal classification.

I. Introduction

In the latest years, the use of wireless communication systems has increased rapidly and the security of these systems has become more important recently. Therefore, wireless communication security should be improved. There are many software improvements, but these improvements are not enough for completely secure systems. Radio frequency fingerprinting (RFF) is one of the communication networks security techniques based on identification of the imperfections caused by a manufacturing process as unique features of the transceivers [1-3]. The applications of RFF technique lie in safe radio communications, radar systems, military communications confrontation and civilian radio monitoring. The RFF technique has also been applied to enhance the security of wireless network, such as Wi-Fi networks, VHF radio networks and cellular networks and so on. Instantaneous amplitude, phase and frequency are the main elements for RFF process. Statistical features such as variance, kurtosis and skewness are the features that can be extracted from these main elements of transient signal [4, 5]. Features can be obtained from different parts of the captured radio signals, such as turn-on transients, RF burst signals, preambles, etc. RFF techniques can be split into two groups, namely transient based and modulation based RFF. Transient based technique extracts features from transient signals that exist at the beginning of the transmissions. Otherwise, modulation based techniques concentrate on modulator imperfections of the transceiver. The importance of security for communication networks, prompted us to find out how can we develop and improve RFF as wireless networks security. When the emitting signals of wireless devices are noisy, it is difficult to distinguish the different wireless transmitters of these signals. Therefore, we need to improve the signal processing functions on the captured signal before extracting the features. In this paper, we propose to use variational mode decomposition (VMD) in RFF of Bluetooth signals before transient detection. VMD is a new modal decomposition method, proposed in latest years as denoising method [6, 7]. VMD is a very good robustness against the effect of background noise on the accuracy of signal processing.

II. Variational Mode Decomposition

Variational mode decomposition (VMD) is a decomposition method [6]. The objective of VMD is to decompose the input signal into different discrete modes $u_z\ (z = 1,2,...,Z)$, where Z is the number of the decomposed modes. Each discrete mode has limited bandwidth in the spectral domain, where each mode is supposed to be compacting around a centre frequency $\omega_z\ (z = 1,2,...,Z)$. The reconstructed signal is equal to the sum of the decomposition modes. The input signal bandwidth is computed as follows: (1) the analytic signal for each mode u_z is constructed by means of the Hilbert transform to achieve a one side frequency spectrum; (2) by using demodulation method, the frequency spectrum of each mode is shifted to estimated baseband; (3) Finally, the squared L^2 norm of the gradient of the demodulated signal is applied to estimate the bandwidth of each mode [6, 8]. The measure of bandwidth of the modes can be calculated by the integral of the square of the time derivative of the shifted, demodulated mode.

$$i.e. \Delta w_z = \int |\partial_t(u_z^D)|^2\ dt \quad = \|\partial_t(u_z^D)\|^2 \tag{1}$$

Received-August 2020, Accepted-October 2020, and Published-December 2020

Therefore, the required minimization problem is

$$\min_{\{u_z\},\{\omega_z\}} \left\{ \sum_z \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_z(t) \right] e^{-j\omega_z t} \right\|_2^2 \right\} \quad (2)$$

To transform the optimization problem into an unconstrained problem form, the quadratic penalty function (α) and Lagrange multiplier (λ) are introduced.

$$L(u_z,\omega_z,\lambda) = \alpha \sum_z \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_z(t) \right] e^{-j\omega_z t} \right\|_2^2 + \left\| f(t) - \sum_i u_i(t) + \frac{\lambda(t)}{2} \right\|_2^2 - \left\| \frac{\lambda(t)^2}{4} \right\|_2^2 \quad (3)$$

Minimization with respect to u_z using Alternate Direction Method of Multipliers (ADMM) algorithm, the sub-problem is rewritten as the following:

$$u_z^{n+1} = \arg\min \left\{ \alpha \sum_z \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_z(t) \right] e^{-j\omega_z t} \right\|_2^2 + \left\| f(t) - \sum_i u_i(t) + \frac{\lambda(t)}{2} \right\|_2^2 \right\} \quad (4)$$

This problem can be solved in spectral domain and both terms written as half-space integrals over the non-negative frequencies:

$$\hat{u}_z^{n+1} = \arg\min \left\{ \int_0^\infty 4\alpha(\omega - \omega_z)^2 |\hat{u}_z(\omega)|^2 + 2\left| \hat{f}(\omega) - \sum_i \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2} \right|^2 d\omega \right\} \quad (5)$$

The solution of this quadratic optimization problem is as following:

$$\hat{u}_z^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i\neq z} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1+2\alpha(\omega-\omega_z)^2}, \quad (6)$$

The centre frequencies ω_z appear only in the first term. This quadratic problem is solved as:

$$\omega_z^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_z(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_z(\omega)|^2 d\omega}, \quad (7)$$

VMD method depends on a number of input parameters like; the balancing parameter (α), the time step (τ), the number of modes (Z), the tolerance of convergence (ε), and the initialization of centre frequencies (ω_z^0), z=1,....,Z. In our proposed method, the values of VMD input parameters namely ε, τ, ω_z^0, α, and Z, have been fixed to the values 10^{-7}, 0, 0, 200 and 3, respectively [6-8].

III. RF FINGERPRINT IMPLEMENTATION

In this work, we present our VMD based method for RF fingerprinting of BT devices. The method has three major stages: Recorded BT signal decomposition using VMD; feature extraction; and BT devices identification using a linear support vector machine (LSVM) classifier. In the first stage, the received BT signal is decomposed into number of modes. In this work, three numbers of modes are chosen for decomposing the BT signal after experimental analysis. In the second stage, different statistical features are extracted from the transient region which detected from the reconstruction BT signal. In the final stage, these extracted features from the transients of different BT devices are fed to a linear support vector machine (LSVM) classifier for BT devices identification. These stages are detailed in the following subsections.

A. Data Acquisition

Bluetooth signals were captured in the laboratory [8, 9]. The laboratory is located in an isolated location (underground) where no other equipment is operating, as any variation in temperature or humidity status may impact the accuracy of the identification. Temperature and humidity are kept constant in the laboratory. Figure1 illustrates the signal collection system. The data is collected and stored in the laboratory. The carrier frequency bandwidth of the Bluetooth signal ranges from 2400 MHz to 2483.5 MHz is received by a typical modem antenna and then applied into an oscilloscope (4GHz). BT signals were recorded using a sampling frequency of 20 GHz. Bluetooth signals radiated by mobile phones were received in the laboratory by an oscilloscope device of sampling rate (20 GSPS). One hundred and fifty Bluetooth signals were captured for each cell phone. Five brands (Huawe, I-Phone, L-G, Samsung, and Sony) of ten various models were chosen for the testing. Two models were selected for each mobile phone brand, and two mobile phones with various serial numbers were selected for each model. Therefore, a data set of twenty mobile phones containing one hundred and fifty transient signals was formed for each mobile phone.

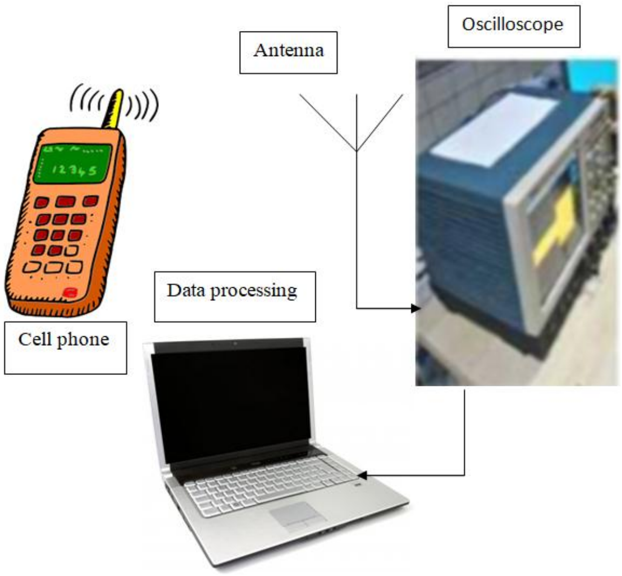

Fig. 1. Signal acquisition systems

B. Signal Decomposition and Transient Detection

Transient signals have unique characteristics which can be extracted for identification of the BT devices. Therefore an accurate detection of transient signal increases the performance of identification systems. Detection of transient signal can be obtained by means of signal characteristics such as frequency, amplitude, phase, and energy [10]. The collected BT signals are de-noised by using VMD method, to remove unwanted components and noise, before the transient detection. The data is presented in three ranges of SNR, low range (5 dB-10 dB), moderate range (10 dB-15 dB) and high range (15 dB-25 dB). Before transient detection, each BT signal x(n) is decomposed by VMD and the transients are detected from the reconstructed BT signal g(n). VMD decomposes the received BT signal x(n) into a specified number of modes

Received-August 2020, Accepted-October 2020, and Published-December 2020

with good spectral separation and less computational complexity as compared to other methods [6]. Furthermore, it provides both time and spectral modes with their centre frequency. Therefore in this work, we decompose the received BT signal x(n) into three number of modes empirically. Let, Z be the number of decomposed modes then the output of VMD can be written as g(n) = $\sum_1^Z u_z$, z=1,2,3....,Z, where, u_z is the temporal mode. The transients are detected from the reconstruction BT signal g(n) which equal to the sum of the modes. The most popular transient detection method was based on the energy envelope of signal. This method was developed in [10] and was used for transient detection in our study. Figure 2 shows a sample outcome of the transient detection technique where the start point and the end point of the transient are indicated.

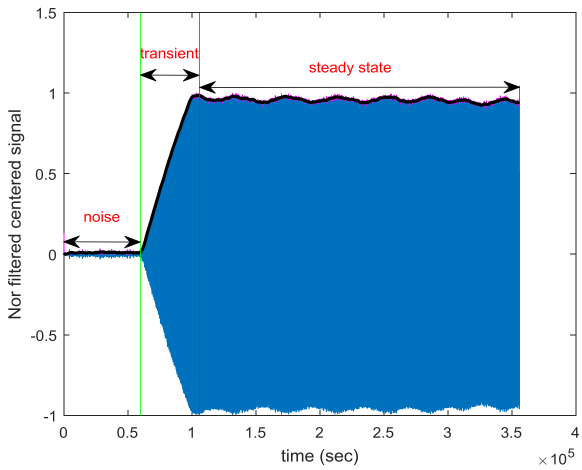

Fig. 2. Decomposed BT signal and transient detection

C. Features extraction

RF fingerprint of a signal can be obtained from its instantaneous amplitude $a(n)$, instantaneous frequency $f(n)$ and phase $\emptyset(n)$. To calculate $a(n)$, $\emptyset(n)$ and $f(n)$, the real valued transient y(n) is converted to complex I and Q data using a Hilbert transform [1 ,2].

We can define the complex signal $\hat{y}(n)$ as

$$\hat{y}(n) = y(n) + jH\{y(n)\} = A(n)e^{j\theta(n)} \tag{8}$$

$$\hat{y}(n) = y(n) + j(h(n) * y(n)) \tag{9}$$

The complex time domain signal $\hat{y}(n)$ can be written in the form:

$$\hat{y}(n) = \hat{y}_I(n) + \hat{y}_Q(n) \tag{10}$$

Has instantaneous amplitude, $a(n)$, phase, $\emptyset(n)$, and frequency, $f(n)$, given by

$$a(n) = \sqrt{\left(\hat{y}_I(n)\right)^2 + \left(\hat{y}_Q(n)\right)^2} \tag{11}$$

$$\emptyset(n) = tan^{-1}\left[\frac{\hat{y}_Q(n)}{\hat{y}_I(n)}\right], \tag{12}$$

$$f(n) = \frac{\emptyset(n) - \emptyset(n-1)}{2\pi\Delta(n)} \tag{13}$$

The instantaneous amplitude, frequency response and phase are simply centred using

$$a_c(n) = a(n) - \mu_a \tag{14}$$

$$f_c(n) = f(n) - \mu_f \tag{15}$$

$$\emptyset_{cnl}(n) = \emptyset_{nl}(n) - \mu_{\emptyset nl} \tag{16}$$

$$\emptyset_{nl}(n) = \emptyset(n) - 2\pi\mu_f(n)\Delta t \tag{17}$$

Where n=1, 2, 3, N is the total number of samples in the signal and μ_a, μ_f are amplitude and frequency means. $\emptyset_{nl}(n)$ is the resultant non-linear phase response.

Statistical features are generated as following:

Variance $$\sigma_v^2 = \frac{1}{N_v}\sum_{k=1}^{N_v}[v(k) - \bar{v}]^2 \tag{18}$$

Skewness $$\gamma_v = \frac{1}{\sigma_v^3 N_v}\sum_{k=1}^{N_v}[v(k) - \bar{v}]^3 \tag{19}$$

Kurtosis $$k_v = \frac{1}{\sigma_v^4 N_v}\sum_{k=1}^{N_v}[v(k) - \bar{v}]^4 \tag{20}$$

Where $v(k)$ represents any of the sequences given in equations (14), (15), and (16) after they normalized. The number of features can be created by using these sequences with these three high order statistics. Characteristics of the descriptive statistics can be visualized by means of boxplot of data. The edges of the boxes correspond to the 25th percentile, also called the first quartile (Q1), and the 75th percentile, also called the third Quartile (Q3), values. The notches of the boxes indicate the median which is also called the second quartile (Q2) values. The lengths of a boxes give interquartile range (IQR=Q3-Q1) values. The edges of dashed lines represent the most extreme data points. Figure 3 shows the variance (VAR) of the transients' instantaneous phases (INS.PH). As seen from the figure, there are statistically significant differences between some of the statistics of different classes while some of them are close to overlap. The feature is completely separable, if the box values, including the whiskers and outlier values, of a certain class are separated from the boxes' values of other considered classes and is considered to be a nearly separable feature, if only the whiskers or outliers are inseparable.

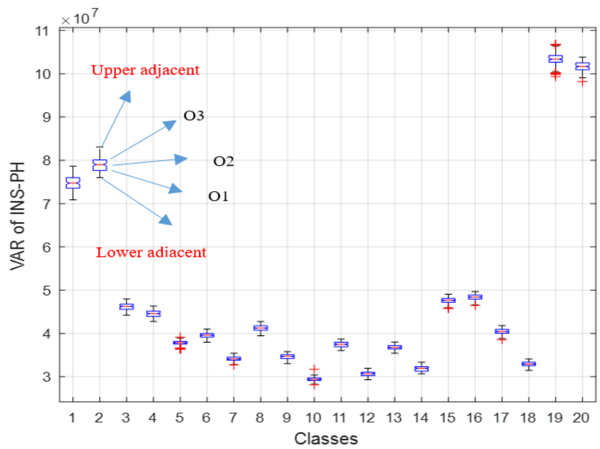

Fig. 3. VAR of INS. PH for 20 mobile phones' transients

D. Classification

When a set of feature vectors have been created, classification can take place. In order to make classification, the feature vectors have to be divided into training data and test data for each Bluetooth device [4, 8]. The training feature vectors serve to establish a relationship between the feature vectors and the BT device from which they are formed. This is done by presenting the classifier

Received-August 2020, Accepted-October 2020, and Published-December 2020

with a feature vector and an associated BT device label for all feature vectors in the training group. Then, the test data of the feature vectors is used to estimate the performance of the classifier. In this stage, each feature vector in the test data is fed to the classifier without a label, and the classifier returns the label of the BT device that it thinks will most correspond to the feature vector.

IV. Results and Discussion

In this work, nine extracted features represent the RF fingerprints of 20 classes, each with 150 records (transients). Each extracted feature data of each class is segmented into two data groups named as: training set and testing set, with 40% and 60% of the total data, respectively. For each set, a linear support vector machine (LSVM) is applied. After the training process, the considered LSVM classifier is applied to the test data which represents 60% of the total data. Figure 4 shows the confusion matrix for 20 Bluetooth devices (mobile phones) under low SNR range (5dB-10dB). The first column shows the actual class and the remaining columns show the predicted class. For example, when class 9 is tested for classification in low SNR, identified 93.3% is correctly as class 9 and 6.7 %, is miss-classified as class 7. However, comparing with the classification in moderate SNR and high SNR, class 9 is identified 96.7 % and 100% correct right, while 3.3 %, is miss- classified as class 7 in moderate SNR. The testing shows different performances for three different SNR ranges of the input data as illustrated in Table 1. It can be seen from the table that the classification accuracy of low SNR, moderate SNR and high SNR are 97.9%, 98.7% and 99.0%. While 38, 21 and 18 records (transients) are miss-classified out of 1800 records (transients) in the testing confusion matrix for low SNR, moderate SNR and high SNR, respectively.

TABLE I. Classification Performance

SNR (dB)	*low (5-10) dB*	*moderate (10-15) dB*	*high (15-25) dB*
VMD based on recorded BT signal method	*97.9%*	*98.7%*	*99.0%*
VMD based on transient region method [8]	*97.4%*	*98.5%*	*98.8%*

Predicted Class

Target Class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	90 100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
2	1 1.1%	89 98.9%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
3	0%	0%	88 97.8%	2 2.2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
4	0%	0%	1 1.1%	89 98.9%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
5	0%	0%	0%	0%	90 100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
6	0%	0%	0%	0%	0%	90 100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
7	0%	0%	0%	0%	0%	0%	90 100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
8	0%	0%	0%	0%	0%	0%	0%	90 100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
9	0%	0%	0%	0%	0%	0%	6 6.7%	0%	84 93.3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
10	0%	0%	0%	0%	0%	0%	0%	0%	0%	83 92.2%	0%	7 7.8%	0%	0%	0%	0%	0%	0%	0%	0%
11	0%	0%	0%	0%	3 3.3%	0%	0%	0%	0%	0%	84 93.3%	0%	3 3.3%	0%	0%	0%	0%	0%	0%	0%
12	0%	0%	0%	0%	0%	0%	0%	0%	0%	1 1.1%	0%	86 95.6%	0%	3 3.3%	0%	0%	0%	0%	0%	0%
13	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	5 5.6%	0%	85 94.4%	0%	0%	0%	0%	0%	0%	0%
14	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	90 100%	0%	0%	0%	0%	0%	0%
15	0%	0%	2 2.2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	88 97.8%	0%	0%	0%	0%	0%
16	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1 1.1%	89 98.9%	0%	0%	0%	0%
17	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	90 100%	0%	0%	0%
18	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	90 100%	0%	0%
19	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	88 97.8%	2 2.2%
20	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1 1.1%	89 98.9%

Fig. 4. Confusion matrix for low SNR values

Received-August 2020, Accepted-October 2020, and Published-December 2020

V. Conclusions

In this paper, we propose an identification method of Bluetooth devices based on decomposition of the recorded BT signal using VMD. The method has three major stages: In the first stage, VMD is used to decompose the BT signal into a series of band- limited modes. Then the selected modes are added to reconstruct the original BT signal. In the second stage, different statistical features are extracted from the transient instantaneous responses such as variance, kurtosis and skewness. In the final stage, a LSVM classifier is used to identify twenty BT devices (mobile phones) under different SNR condition. In order to test the performance at different SNR levels, the signals with SNR values of (5-10dB), (10-15dB) and (15-25dB) are used to test. We evaluated the performance of the proposed method in 20 mobile phones with 150 transients per device by using the correct classification percentage (cc%). To demonstrate the performance of our proposed method, we compared it with the existing method that based on transient decomposition in terms of (cc%). Results show that the proposed VMD identification method based on full signal (recorded signal) decomposition outperforms the existing method at different ranges of SNR. In future direction, the proposed method can be applicable for any wireless network security purpose, such as Wi-Fi, Wi-Max, radar system, and so on.

References

[1] M. Köse, S. Taşcioğlu, and Z. Telatar, '''RF fingerprinting of IoT devices based on transient energy spectrum,'' *IEEE Access*, vol. 7, pp. 18715–18726, 2019.

[2] A. M. Ali, E. Uzundurukan, and A. Kara, "Assessment of features and classifiers for Bluetooth RF fingerprinting," *IEEE Access*, vol. 7, pp. 50524-50535, 2019.

[3] E. Uzundurukan, Y. Dalveren, and A. Kara, ''A Database for the Radio Frequency Fingerprinting of Bluetooth Devices," *Data*, vol. 5, no.55, pp. 1-11, June 2020.

[4] A. Aghnaiya, Y. Dalveren, and A. Kara, ''on the performance of variational mode decomposition-based radio frequency fingerprinting of Bluetooth devices,'' Sensors, vol. 20, no. 6, p. 1704, Mar. 2020..

[5] R. M. Ferre, W. Wang, A. S. Abia, and E. S. Lohan, ''Identifying GNSS Signals Based on Their Radio Frequency (RF) Features—A Dataset with GNSS Raw Signals Based on Roof Antennas and Spectracom Generator," *Data*, vol. 5, no.18, pp. 1-13, February 2020.

[6] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transaction Signal Process*, vol.62, no.3, pp. 531-534, 2014.

[7] S. A. Abufana, Y. Dalveren, A. Aghnaiya, and A. Kara, ''Variational Mode Decomposition-Based Threat Classification for Fiber Optic Distributed Acoustic Sensing," *IEEE Access*, vol. 8, pp. 100152-100158, May.2020.

[8] A. Aghnaiya, A. M. Ali, and A. Kara, ''Variational mode decomposition-based radio frequency fingerprinting of Bluetooth devices,'' *IEEE Access*, vol. 7, pp. 144054–144058, October 2019.

[9] I. S. Mohamed, Y. Dalveren, and A. Kara, ''Performance Assessment of Transient Signal Detection Methods and Superiority of Energy Criterion (EC) Method," *IEEE Access*, vol. 8, pp. 115613-115620, June.2020.

[10] A. M. Ali, E. Uzundurukan, and A. Kara, "Improvements on transient signal detection for RF fingerprinting," *Signal Processing and Communications Applications Conference*, Antalya, Turkey, pp. 1-4,2017.

Received-August 2020, Accepted-October 2020, and Published-December 2020