A. ALI

ATILIM UNIVERSITY 2019

ASSESSMENT OF FEATURES AND CLASSIFIERS FOR BLUETOOTH RF FINGERPRINTING

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
ATILIM UNIVERSITY

AYSHA ALI

DOCTOR OF PHILOSOPHY THESIS
IN
MODELING AND DESIGN OF ENGINEERING SYSTEMS (MAIN FIELD OF STUDY: ELECTRICAL ELECTRONICS ENGINEERING)

MAY 2019

ASSESSMENT OF FEATURES AND CLASSIFIERS FOR BLUETOOTH RF FINGERPRINTING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
ATILIM UNIVERSITY

BY

AYSHA ALI

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
MODELING AND DESIGN OF ENGINEERING SYSTEMS
(MAIN FIELD OF STUDY: ELECTRICAL ELECTRONICS ENGINEERING)

MAY 2019

Approval of the Graduate School of Natural and Applied Sciences, Atilim University.

_____

Prof. Dr. Ali Kara
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of **Doctor of Philosophy in Electrical and Electronics Engineering, Atilim University.**

_____

Prof. Dr. Ender Keskinkılıç
Head of Department

This is to certify that we have read the thesis ASSESSMENT OF FEATURES AND CLASSIFIERS FOR BLUETOOTH RF FINGERPRINTING submitted by AYSHA ALI and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Doctor of Philosophy.

_____

Prof. Dr. Ali Kara
Supervisor

Examining Committee Members:

Prof. Dr. Şimşek Demir
Electric and Electronics Eng. Department, METU _____

Prof. Dr. Ali Kara
Electric and Electronics Eng. Department,
Atilim University _____

Assoc. Prof. Dr. Murat Hüsnü Sazlı
Electric and Electronics Eng. Department,
Ankara University _____

Asst. Prof. Dr. Ertan Zencir
Electric and Electronics Eng. Department,
Turkish Aeronautical Association University _____

Asst. Prof. Dr. Yaser Dalveren
Civil Aviation Department, Atilim University _____

Date: *3 May 2019*

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Aysha, Ali :

Signature :

ABSTRACT

ASSESSMENT OF FEATURES AND CLASSIFIERS FOR BLUETOOTH RF FINGERPRINTING

Ali, Aysha
Phd., Modelıng And Desıgn Of Engıneerıng Systems

Supervisor : Prof. Dr. Ali Kara

May 2019, 99 pages

In this thesis, we introduced a novel technique to enhance the security at physical layer of wireless networks. This is based on the use of radio freqency (RF) fingerprinting for Bluetooth (BT) signals. BT signal records are acquired from twenty different cell phone brands, models, and serial numbers. One hundred fifty records are collected from each device. For the first time, Hilbert Huang Transform (HHT) are used for the BT device identification with such huge data set. By means of the signals' energy envelopes with some improvements, the transient signals are detected accurately. Through the Empirical mode decomposition (EMD) and Hilbert Transform (HT), the HHT is implemented to obtain Time Frequency Energy Distributions (TFED) of the detected transients. Thirteen features are extracted from the signals' transients and their TFEDs. The extracted features are pre-processed to enhance their usability. Different classifiers are employed with the extracted features for device identification, and comparative analysis of the classifiers is also provided. The classifier performance is examined for different SNR levels from 8 dB to 35+ dB . The identification performance demonstrates the feasibility of the method.

Keywords: Bluetooth, Radio frequency, Transient detection, Hilbert Huang Transform, Classification.

ÖZ

BLUETOOTH SİNYALLERİNİN RF PARMAK İZİ YÖNTEMİ İLE SINIFLANDIRILMASINDA ÖZNİTELİKLER VE SINIFLANDIRICILARIN DEĞERLENDİRİLMESİ

Ali, Aysha
Doktora, Mühendislik Sistemlerinin Modellenmesi Ve Tasarımı
Tez Yöneticisi : Prof. Dr. Ali Kara

Mayıs 2019, 99 sayfa

Bu tez çalışması, kablosuz ağların güvenliği için fiziksel katmanda özgün yöntemler geliştirilmesi üzerinedir. Bunun için, RF parmak izi kullanılarak, Bluetooth (BT) sinyalleri üzerinde çalışılmıştır. RF parmak izi tespiti için kapsamlı BT sinyalleri kullanılmıştır. Bu kapsamda, 20 farklı marka, model ve seride BT cihazına yönelik kayıtlar toplanmıştır. Her bir cihazdan da 150 sinyal kaydı alınmıştır. Bu çalışma ile Hilbert-Huang Dönüşümü (HHT) ilk defa BT cihaz kimliklendirmesinde kullanılmıştır. Amprik Kip Ayrıştırma (EMD) ve Hilbert Dönüşümü kullanımı ile HHT tekniği, zaman-frekans-enerji dağılımları üzerinde çalışma imkanı vermektedir. Sinyal enerji zarfı kullanılması suretiyle geçici rejim sinyalleri bazı iyileştirmeler ile tespit edilmektedir. Geçici rejim sinyalleri ile zaman-frekans-enerji dağımları üzerinden toplam 13 farklı öznitelik çıkarılmaktadır. Öznitelikler, kullanılabilirlik açısından ön işleme tabi tutulmaktadırlar. Ardından aynı veri seti ve öznitelikler üzerinde farklı sınıflandırıcılar çalıştırılarak, sınıflandırıcıların başarım analizi de ilk defa bu çalışmada sunulmaktadır. Sınıflandırıcı başarım analizleri 8 dB ile 30+ dB arasında farklı sinyal-gürültü oranlarında yapılmaktadır. Sınıflandırma başarım sonuçları yöntemin kullanılabilirliğini göstermektedir.

Anahtar Kelimeler: Bluetooth (BT), RF parmak izi, Geçici Rejim Sinyali Tespiti,Hilbert-Huang Dönüşümü, Sınıflandırma.

To my family

ACKNOWLEDGMENTS

Praise to Allah, Lord of the worlds. I would like to express my special appreciation and thanks to my advisor Professor Dr. Ali Kara. your encouragement, helpful advice, unparalleled support, and unwavering guidance during my research study inspired me to work hard, confident, and continual.

Furthermore, I would also like to extend my deepest gratitude to my committee members, Assoc. Prof. Dr. Murat Hüsnü Sazlı and Asst. Prof. Ertan Zencir ,and Asst. Prof. Dr. Harun Taha Hayvacı for offering constructive criticism, insightful suggestions, and patience that cannot be underestimated. My deep appreciation goes out to the members of my examination committee Prof. Dr. Şimşek Demir and Asst. Dr. Yaser Dalveren.

Finally, no word can express my thanks to my hausband, who has made sacrifices on my behalf. I am really deeply indebted to him for his supporting.

TABLE OF CONTENTS

ABSTRACT iii
ÖZ iv
DEDICATION v
ACKNOWLEDGMENTS vi
TABLE OF CONTENTS vii
LIST OF TABLES ix
LIST OF FIGURES x
LIST OF SYMBOLS/ABBREVIATIONS xiii
CHAPTER
1. INTRODUCTION 1
2. BLUETOOTH RADIO SYSTEM 7
2.1 The different layers of Bluetooth protocol 9
2.2 Some advantages of Bluetooth 11
2.3 Bluetooth radio spectrum 12
2.3.1 Frequency hopping 12
2.3.2 Frequency bands and channel arrangement 13
2.4 Bluetooth packets 14
2.5 The modulation scheme 15
2.6 Network topology 16
2.6.1 Point to point and point to multipoint piconets 16
2.6.2 Scatternets 17
3. HILBERT HUANG TRANSFORM 19
3.1 Empirical Mode Decomposition 21
3.1.1 The steps of EMD 21
3.2 The Hilbert spectral analysis 23
3.2.1 Signal characteristics 24
3.3 Hilbert Spectrum 25

4. TRANSIENT SIGNAL DETECTION 29
4.1 Start transient detection 30
4.1.1 Threshold detection approach 30
4.1.2 Bayesian change point detection 32
4.1.3 Transient detection based on phase characteristic 34
4.2 Detection of transient start and end points based on improved methods 36
4.2.1 Improved Bayesian change point detection 36
4.2.2 Improved Phase based method 38
5. IMPLEMENTATION 41
5.1 Data acquısition 41
5.2 Detection of transient start and end points based on energy envelop . 45
5.3 Obtaining the Hilbert spectrum 48
5.4 Extracting and smoothing features 55
5.5 Classification based on high SNR level 66
6. NOISE PERFORMANCE ANALYSIS 77
6.1 Unsmoothed features 80
6.1.1 Classifiers training for unsmoothed features 80
6.1.2 Classifiers testing for unsmoothed features 84
6.2 Classifiers testing for smoothed features 87
7. CONCLUSIONS 92
REFERENCES 95

LIST OF TABLES

TABLES

Table 5.1 The classes 42

Table 5.2 A sample of some different features of four different cell phone transient signals..... 55

Table 5.3 RF fingerprints .....51 Table 5.3 RF fingerprints 56

Table 5.4 Classifiers' accuracies based on 40%-train 60%-test unsmoothed RF fingerprint..... 68

Table 5.5 Classifiers' accuracies based on 20%-train 80%-test unsmoothed RF fingerprint..... 69

Table 5.6 Classifiers' accuracies based on 40%-train 60%-test smoothed RF fingerprint..... 74

Table 5.7 Classifiers' performances at high SNR (30-35dB) 76

Table 6.1 Training accuracies of classifiers for unsmoothed features at different SNR levels..... 83

Table 6.2 Areas under ROC curves with different levels of SNR for different classifiers based on unsmoothed features 84

Table 6.3 Classifiers' test performances with different levels of SNR (un-smoothed features)..... 87

Table 6.4 Classifiers' test performances with different levels of SNR (smoothed features)..... 90

LIST OF FIGURES

FIGURES

Figure 2.1 Buetooth operating spectrum 9
Figure 2.2 Layers of Bluetooth protocol 11
Figure 2.3 Transmitting Bluetooth wireless packaets over multiple channels..... 13
Figure 2.4 Bluetooth frequency bands and RF channels 14
Figure 2.5 Bluetooth wireless packet format 15
Figure 2.6 Bluetooth network topology 18
Figure 3.1 A flowchart of generating Hilbert spectrum 28
Figure 4.1 A flowchart of of The improved methodes..... 37
Figure 4.2 Flowchart of the improved phase based method 39
Figure 4.3 Comparison between the three different detection techniques 40
Figure 5.1 Illustration of signal aquisition system 43
Figure 5.2 A typical form of the collected BT signal 43
Figure 5.3 Filtered BT signals..... 44
Figure 5.4 A typical form of filtered normalized BT signal 45
Figure 5.5 The detected transients of BT signals 48
Figure 5.6 The 1st IMFs of transient cell phone signals 49
Figure 5.7 The 2nd through 5th IMFs of transient signals 50
Figure 5.8 The TFEDs plane of transient signals..... 51
Figure 5.9 The zoomed TFEDs plane of transient signals 52
Figure 5.10 Three dimensional TFEDs of transient signals..... 53
Figure 5.11a The summation of HS energy along time axis..... 54
Figure 5.11b The summation of HS energy along frequency axis..... 54
Figure 5.12 The box plot criteria..... 57
Figure 5.13 A conventional plot of STD of trans. ins. phases of 20 cell phones..... 58
Figure 5.14 The box plot of STD of trans. ins. phases of 20 cell phones 59
Figure 5.15 Zooming of the STD of transients' instantaneous phases 60

Figure 5.16 Box plot of zoomed STD of transients' instantaneous phases 61
Figure 5.17 Smoothing of the STD of transients' instantaneous phases..... 62
Figure 5.18 Box plot of smoothed STD of transients' instantaneous phases 62
Figure 5.19 Unsmoothed vs. smoothed STD of six classes trans. ins. phase 63
Figure 5.20 Box plot of smoothed entropy of transients' instantaneous phase 64
Figure 5.21 Features distribution of class-13 & class-14..... 65
Figure 5.22 Features distribution of class-8 & class-9..... 66
Figure 5.23 Scatter plot of unsmoothed features f1 vs . f2 70
Figure 5.24 Scatter plot of unsmoothed features f1 vs. f8 70
Figure 5.25 Scatter plot of smoothed features f1 vs. f2 71
Figure 5.26 Scatter plot of smoothed features f1 vs. f8 71
Figure 5.27 Complex tree classifier confusion matrix based on 40%-train of un-smoothed RF fingerprint 72
Figure 5.28 Complex tree classifier confusion matrix based on 40%-train of smoothed RF fingerprint 73
Figure 5.29 Test confusion matrix for 30-35 high SNR un-smoothed transients via L-SVM 75
Figure 5.30 Test confusion matrix for 30-35 high SNR smoothed transients via L-SVM 76
Figure 6.1 A recorded channel noise and associated transient signal 78
Figure 6.2 The Histogram of SNR level values of 3000 noisy transient records..... 78
Figure 6.3 Block diagram of process of adding noise to transient..... 79
Figure 6.4 Train confusion matrix for 12-15 SNR unsmoothed transients via LDA 81
Figure 6.5 Train confusion matrix for 12-15 SNR unsmoothed transients via L-SVM 81
Figure 6.6 Train confusion matrix for 12-15 SNR unsmoothed transients via Complex Tree..... 82
Figure 6.7 The ROC and AUC for 12-15 SNR un-smoothed arbitrary transients... 83
Figure 6.8 Test confusion matrix for 12-15 SNR un-smoothed transients via LDA 85
Figure 6.9 Test confusion matrix for 12-15 SNR un-smoothed transients via

L-SVM 86
Figure 6.10 Test confusion matrix for 12-15 SNR un-smoothed transients via Complex Tree 86
Figure 6.11 Test confusion matrix for 12-15 SNR smoothed transients via LDA .. 88
Figure 6.12 Test confusion matrix for 12-15 SNR smoothed transients via L-SVM 89
Figure 6.13 Test confusion matrix for 12-15 SNR smoothed transients via Complex Tree 90

LIST OF SYMBOLS/ABBREVIATIONS

ACK	Acknowl- edgement
ACL	Asynchronous Connectionless Links
ADC	Analog to Digital Converter
AV	Unwrapped Phase Vector
BPF	Band Pass Filter
BT	Bluetooth
C-Tree	Complex Tree
CPV	Cauchy Principle Value
DHT	Discrete Hilbert Transform
DPSK	Differential Phase Shift Keying
DQPSK	Differential Quadrature Phase Shift Keying
EDR	Enhanced Data Rate
EMD	Empirical Mode Decomposition
FH	Frequency Hopping
FT	Fractal Trajectory
GFSK	Gaussian Frequency Shift Keying
HCI	Host Controller Interface
HHT	Hilbert Huang Transform
HS	Hilbert Spectrum
HSA	Hilbert Spectrum Analysis
HT	Hilbert Transform
IBM	Intrenet Business Machine
IF	Instantaneous Frequency
IMF	Intrinsic Mode Function
ISM	Industrial, Scientific, and Medical
IQR	Inter Quartile Range
KNN	Kth Nearest Neighbors

LANs	Local Area Networks
LC	Link Controller
LDA	Linear Discriminant Analysis
LM	Link Management
LMP	Link Manager Protocol
L-SVM	Linear Support Vector Machine
MAC	Media Access Control
NACK	Negative Acknowledgement
NASA	National Aeronautics and Space Administration
OSI	Open Systems Interconnection
PAN	Personal Area Network
PHY	Physical Layer
PSK	Phase Shift Keying
RF	Radio Frquency
PSK	Phase Shift Keying
SCO	Synchronous Connection-Oriented Link
SDP	Service Discovery Protocol
SEI	Spesific Emitter Identification
SNR	Signal to Noise Ratio
STD	Standard Deviation
STFT	Short Time Fourier Transform
SVM	Support Vector Machine
TFED	Time Frequency Energy Distribution
TFR	Time Frequency Representation
WPAN	Wireless Personal Area Networks
WT	Wavelet Transform
WV	Wigner Ville
WVD	Wigner Ville Distribution

CHAPTER 1

INTRODUCTION

Recently, wireless network has become an essential part of our daily life, it is vastly used in both civilian and military situations. Since wireless networking plays an extremely important role in transmission data [1], security has become a fundamental attention in order to protect communication networks from some threats such as signal interception, spoofing or jamming. Those threats considered to be major attacks at the physical layer of a wireless network. However, security of transmited information against such attacks still a challenging issue [1].]. Extensive studies have been conducted to enhance the security at the Open System Interconnection (OSI) of Physical (PHY) layer to protect the confidentiality of wireless networks. Some methods to enhance the security at the physical layer in wireless networks are introduced. Radio Frequency (RF) fingerprinting is considered as an enhancement wireless networks security technique, by which a spesific radio transmitter can be identified [2]. RF fingerprinting is a technique that has been used to identify wireless devices in mobile and wireless networks by capturing unique characteristics. RF Fingerprinting technique is accomplished by detect the transient portion of the transmission signal and extract the features of the signal from that portion.

Once a radio transmitter is activated, the RF signal passes through a transient portion which is attributted to some manufacturing imperfections in the hardware of a device's radio transceiver [3]. Those imperfections are the acquisition characteristics of frequency synthesis systems, modulator subsystems, RF amplifiers, and digital-to-analog converters [2]. The transient signals from each emitter usually have special RF fingerprint for spesific emitter identification SEI. RF fingerprinting is proposed as unique features in the signal transient; therefore, it can be used to identify a SEI

[2]. In [1] several common methods are offered in toturial manner to increase security at the physical layer in wireless networks. The methods are classified into five categories based on their characteristic features. The categories are discussed in terms of two metrics channel capacities, and computational complexities. The five categories are theoretical secure capacity, channel, coding, power, and signal detection approaches. In channel approach the RF fingerprinting has been proposed to enhance the security by exploiting the characteristics of the channel. The RF fingerprinting system proposed by [1, 4] consists of multiple sensor systems to capture and extract RF features of the received signals. This RF fingerprinting system observes the interim evolution of each fingerprint and releases an intrusion warning when a strange fingerprint is detected, thus helping differentiate between an intruder and a legitimate user. The feasibility of implementing impersonation attacks on the transient and modulation-based fingerprinting techniques is studied in [3]. The transient-based techniques observe unique features during the transient stage. It is primarily used for distinguishing classes (model) of wireless devices. Modulation based features technique can identify classes and identical devices as well. [5] proposed a RF fingerprinting technique for appropriate identification of wireless devices in mobile and wireless networks. The procedure is accomplished by capturing and analyzing the initial preamble of the extracted RF signal and then compare it with a set of already known device RF fingerprints. In [2] the problem of identifying an individual node in a wireless network is handled by its own RF footprint. Also [6] illustrated a robust RF fingerprinting procedure that is consistently accurate with high- end and low-cost receivers. The method shew that the properties of the fingerprint based on frequency are more suitable for classifying the device when using a low-cost signal receiver rather than a sophisticated receiver. The research represented an important step toward realizing a practical, low-cost RF fingerprinting solution. In the litrature, RF fingerprinting was applied to many different types of signals by means of different methods. In our work RF fingerprinting technique is demonstrated with Bluetooth (BT) transceivers. BT is a short range, low power communication technology used to establish wireless connections between RF devices. BT radio system works in the unlicensed Industrial,

Scientific, and Medical (ISM) band from 2.4000 to 2.4835 GHz [7, 8, 9, 10]. As wireless data is exposed to spying, there is a need to transmit data in a scure mannar. The security is a crucial issue for wireless networks and BT wireless network [7]. In addition, in ISM band there are other systems also work in that region. So that interference with BT devices is unavoidable because other type of equipment such as microwave oven and LANs use the 2.4 GHz frequency. BT system depends on some techniques in order to keep security of transmitted data over wireless links. In order to address the matter, frequency-hopping spread-spectrum techniques are employed. One of these techniques is Frequency Hopping (FH), By which the BT radio system is switching between wireless channels. BT uses 79 radio frequency channels spaced 1 MHz apart in the frequency range of 2.402 to 2.480 GHz. The rate of hopping is 1600 hops per second. In BT network topology up to eight devices can be connected to each other. The connected BT devices form a small net called piconet. In the piconet one device called a master and the other devices called slaves. The Slaves in a piconet only have links to the master; there are no direct links between slaves in a piconet.

The RF Fingerprinting process consists of four essential stages [11]. The first stage involves the extraction of features from the digital signal such as amplitude, phase , frequency, and energy envelope. In the second stage those features are then used to detect the transient of the signal. Once the transient has been detected, the fingerprint is obtained by generating unique features representing the transient in the third stage. Finally, identifying the emitter device based on the classification of the fingerprint.

To identify specific emitter it is important to detect the signal transient and then extract the RF fingerprinting from the transient. This can be done by means of the signal characteristics such as amplitude, frequency, phase, energy envelope of the signal [11]. Detecting the transient is essentially in RF fingerprinting since the transient signal contains the unique features (which will be analyzed to identify the transmissions signal devices), so the more presise detection, the higher classification result will be.

From literaure, to detect the transient signal the start point of the thansient should be detected first. The detection of the start point of the transient has been done by many approaches such as Threshold detection [13], Bayesian Step Change Detector [13][14][15], Phase Based method detection [13][16], and Energy Envelope [17]. The endpoint dection in some references done by the characteristic of the transient energy trajectory [16], and by the energy envelope [17]. In our work we could detect the transient signal (the start and the end points) by means of the energy envelope technique. The advantage of using the energy envelope is that there is no difficulty to select appropriate threshold values suitable for all BT collected signals. Use of energy envelope of signals is very common technique for the detection of transients, more specifically, the start and end points of transients. However, it is necessary to improve the proposed techniques for both more practical implementations and lower computational loads. The algorithm is based on the use of the local energy envelope maxima. The multiple local energy maximums leads to interference in the transient durations of different classes. Then, the transient duration cannot be used as a robust feature. This is resolved by exploiting the transient features to re-capturing the transient end. Therefore, the energy based transient detection algorithm has been improved by introducing several more stages. In this technique, the features to be extracted in the next stage is also considered earlier in the detection of the transients. While it requires more signals from multiple devices, overall, this makes the process more accurate, and keep the features more robust. Once the transient has been detected , the unique features from the transient will be generated; the unique features represent the fingerprint of the radio frequency signal of the wireless device [14]. In previous researches , the RF fingerprint of transient signal is extracted from the instantaneous amplitude, phase, frequency and energy envelope [16]. In [17] some features are extracted from the energy envelope of the transient signal to create the RF fingerprinting of the emitter. In [16] the features are extracted by means of three techniques that are: energy envelope technique, amplitude–phase–frequency technique, and Time Frequency Energy Distribution (TFED) technique. In [18] the effective RF fingerprint can be formed by using the features of the energy trajectories of the transmitters. In our work, features are extracted from three feature groups namely, the transient signal and its energy envelope, summation of TFED of

the transient signal along time axis, and summation of TFED of the transient signal along frequency axis. These features are dominant in constructing the signal RF fingerprint for each record.

The studied devices are classified by means of different classifiers. The classification is not only done based on different brands devices and same brands different model devices but also based on same brands same model different serial number devices. Three different classifiers are applied to the extracted features namely: decision tree, Support Vector Machine, and Linear Discrimenent Analysis classifiers. The considered classifiers are learned by means of the train data and then implemented to the test data.

Before classification process the features are smoothed. The performance of classifiers are analysed and investigated to show the feasibility of identification procedure after and before smoothing. Smoothing process is used in order to enhance the classifiers performance. This can be done by applying a filtering or smoothing technique to the extracted zoomed features. The parameters or the coefficients of the applied filter must be same for all data of classes considered features, otherwise this will corrupt the whole identification process. The median filter is used to smooth the extracted features. After praparation process we need to investigate the robutness of the features before classification processes.

The robustness of the extracted features are investigated by using the Box plot demonstration and conventional two dimensional Matlab plot. By means of the box plot a clear picture of the interference among classes feature can be seen. Moreover, by using the box plot, the percentage of the data concentration can be recognized. By using the conventional plot we can know the number of interfered records and interfered records' numbers for each class. In order to identify the twenty considered classes, each extracted feature data of each class is divided into two data groups namely, training set and testing set. The purpose of splitting data set into training and test data is to see the realistic performance of the classifiers. The two groups of data are divided into 40% and 60% of the total data, train and test set respectively for smoothed and unsmoothed features. And then we used 20% and 80% for train and test data respectively for unsmoothed features. We introduced the different percent of train data and test data for the smoothed and unsmoothed

features to investigate the effect of the data division size. The average accuracy of classifiers' based on 40%-train 60%-test and based on 20%-train 80%-test is compared.

In order to evaluate the effectiveness of the introduced successive signal processes, the classification performances are examined for different levels of Signal to Noise Ratio (SNR) transient signals. The SNR levels of the transients are reduced by adding the recorded channel noise to the transient signals. The noise portion is weighted and added to the transient signal for each record of each class. The weight factors are varied such that the SNR levels of noisy transients are classified into three levels. The considered features are extracted from smoothed and unsmoothed noisy transint signals. The extracted features data are divided into train data and test data. Some predicting criteria are analysied to predict the classifiers' performances. These criteria include training confusion matrix, the receiver operating characteristics (ROC) curves, and the associated areas under curves (AUC). The test confusion matrices are obtained by applying the introduced calssifiers to the noisy test data.

In chapter 2, some information about BT radio system are inroduced. Chapter 3 explains the theory of HS. Chapter 4 introduces some methods from the litrature to detect the signal transient. In chapter 5, the energy envelope based detection technique is implimenated to the BT signals. Then the HHT is applied to the detected transeints to obtain their TFEDs. Also, features are extracted from the transient signals and their TFEDs to generate RF fingerprints. Different classifiers are applied to the generated RF fingerprints to identify and classify the emitters. In chapter 6, the test performances of the Introduced classifiers are analyzed based on smoothed and unsmoothed noisy transient signals. The considered features are extracted from the noisy transient signals and their TFED. The extracted features data are divided into train data and test data. The classifiers are training based on three levels of SNR data. Then confusion test matrices are generated by separately applying the introduced classifiers to smoothed and unsmoothed noisy test data.

CHAPTER 2

BLUETOOTH RADIO SYSTEM

The prosperity of portable devices such as cell phones, laptops, PADs, and many other portable products is contingent upon persistent success of reduction cost and size of the microelectronics circuit. Alongside of amelioration of mobile technology, there is an increasing request on sharing multimedia and documents files between users within a small area. Connection among these portable devices is called Personal Area Networks (PAN). Many wireless standards have been progressed to enable PAN over short ranges to be connected, Bluetooth (BT) is one of the most successful technologies [7].

The name of BT came after the tenth century king monarch named Harald Blatand Gormsson. Blatand literally meaning BT. He has known for uniting moderen day Denmark and Norway in a very peaceful way [8, 10]. In 1994 the Swedish company Ericson mobile communications was looking into an inexpensive, low power radio interface between mobile phones. At that time if you want to connect two mobile phones together, you needed cables and PC cards. So they started doing researches into this. By 1998 Intel, IBM, Toshiba, Ericson, and Nokia all these different combanies were looking into the same kind of inexpensive, low power radio technology. They have decided to come together to form the BT special interest group (a royality free group). So they have brought this group together to create a standard for short range, low power radio link. It operates on the 2.4 GHz band to transmit packets of information from a master device to its slave devicses. It is known to work in the unlicensed Industrial, Scientific, and Medical ISM band. It uses 79 radio frequency channels each spaced 1 MHz apart. In ISM band there are other systems that also work in that region which includes wireless LANs, baby monitors, microwave ovens, medical equipments, and military radars as shown in

fig. 2.1. Wi-Fi can also be used for example (see fig. 2.1) in some Bluetooth applications such as streaming audio; however, Bluetooth has the ability to use the total spectrum which makes it ideal for robust, short-range wireless transmission in crowded radio environments [8].

BT, a short range and low power communication technology, used to establish wireless connections between RF devices [7, 9]. Mainly, it has been used to establish Wireless Personal Area Networks (WPAN), which is indicated as AD-HOC or peer-to-peer networks. By BT technology there is no need for a cable connection between devices. Devices can be connected up to seven to form a piconet.

Bluetooth is a global wireless communication technology determined under the IEEE 802.15.1 Standard [7]. Three different power classes are defined. It uses nominal antenna power of 0dBm for a short range about 10 meters; that means devices must be within 10 meters to each other to communicate using the Bluetooth standard, and uses 4dBm for an ordinary range about 20 meters. 20dBm is used for a long range about 100 meters.

As wireless data is exposed to spying, the security is a crucial issue for wireless networks and BT wireless network [7]. Bluetooth system depends on some techniques in order to keep security of transmitted data over wireless links. To guarantee robust operation in the interference-dominated ISM band, Bluetooth uses many techniques like spread spectrum, frequency hopping, and full-duplex signal at a nominal rate of 1600 hops per second for basic and boosted data rate transmissions. Data is transmitted up to 3 Mbps [10]. General communications protocol limits the practical data rate to about 721 kbps. When there is an interference during transmission, bit error rate will increase and the achievable data rate will decrease. If transmission runs into interference, it waits for the next frequency hop and re-sends a new frequency.

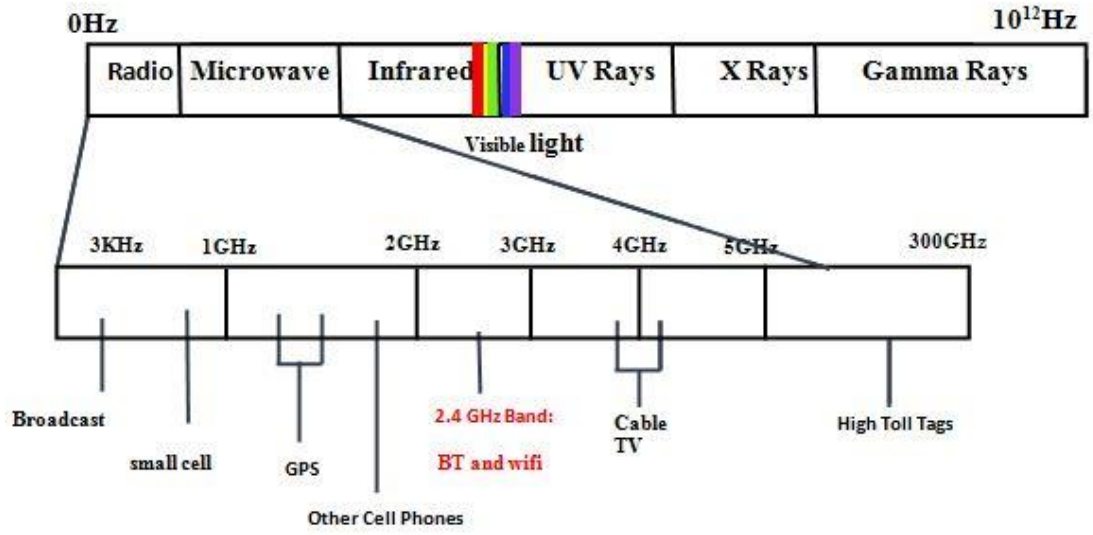


Figure 2. 1 Bluetoooth operating spectrum [8]

2.1. The different layers of Bluetooth protocol

The Bluetooth specification aims to allow devices from different manufacturers to work with one another. Bluetooth does not only just define a radio system, it also defines a software stack to enable applications to find other Bluetooth devices in the area, discover what services they can offer. BT system has different layers protocol. Fig. 2.2 shows the four basic layers and how they interact [10]. The four basic portions are:

I. A radio frequency (RF part) which transmits and receives data and voice
II. A baseband or link control
III. Link management software by which the transmission can be managed
IV. Supporting application software

These basic parts show how Bluetooth works. In Fig. 2.2, The Baseband represents the physical layer of the Bluetooth. It is the hardware that converts received radio signal into a digital form, which can be processed by the host application. The Baseband processes all the functions that performed by the packets. Every packate contains information about its source, what frequency it uses and where it is going to. In addition, it involves information on how the data was packed; how the packets were transmitted and information was used to verify the efficiency of the

transmission. Baseband will check for accuracy when data is received, and extracted from the packet. Baseband also caring for converting data from one to another, e.g., converting from voice to digital data. It will compress the data, put it into packets and take it out of packets, specifying identifiers and error correction information. Then it will reverse the entire process for received data. The baseband function becomes the link controller. The Bluetooth supports two types of link: Synchronous Connection-Oriented Links (SCO) and asynchronous connectionless links (ACL) [9, 10]. SCO which is used mainly for voice communications is a point-to-point link between the master and a single slave on the piconet. ACL which is used for packet data is a point-to-multipoint link between the master and all the slaves on the piconet. Both links support 16 different packets that are used according to the application. The devices in a Bluetooth system can use either link type and can change the link types during a transmission.

The Link Management (LM) software manages the communication between Bluetooth devices. Each BT device has its own link manager. When one of the devices finds out other remote link managers, then it can contact with them through the Link Manager Protocol (LMP) to control link setup, authenticate quality of service, and to encrypt and adjust data rate on link. To perform its service provider role, the LM uses the services of the Link Controller (LC) [10]. LC is a supervisory function that controls all of the Bluetooth baseband functions and supports the link manager. It is responsible for sending and receiving data. It is able to identify the sending device and perform authentication and ciphering functions. Each packet utilizes a single 625 time slot. However, it can be spanned to cover up to five slots. Bluetooth supports an asynchronous data channel, three synchronous voice channels at 64 kbps and simultaneous asynchronous data and synchronous voice channels.

Application software, the last basic section of Bluetooth system, is placed in the device that runs an application over the Bluetooth protocol stack. This software lets the mobile phone to do its job. In order to deal with each Bluetooth device, each of them must have favorable sections in their Bluetooth stack. Thus, all Bluetooth-attested devices must have the basic four parts according to the standard. Besides the four parts that mentioned above, Host Controller Interface (HCI), Logical Link Control and Adaptation Protocol (L2CAP), RFCOMM Protocol and Service

Discovery Protocol (SDP) are needed [10]. The rest parts of the Bluetooth Stack are also very considerable.

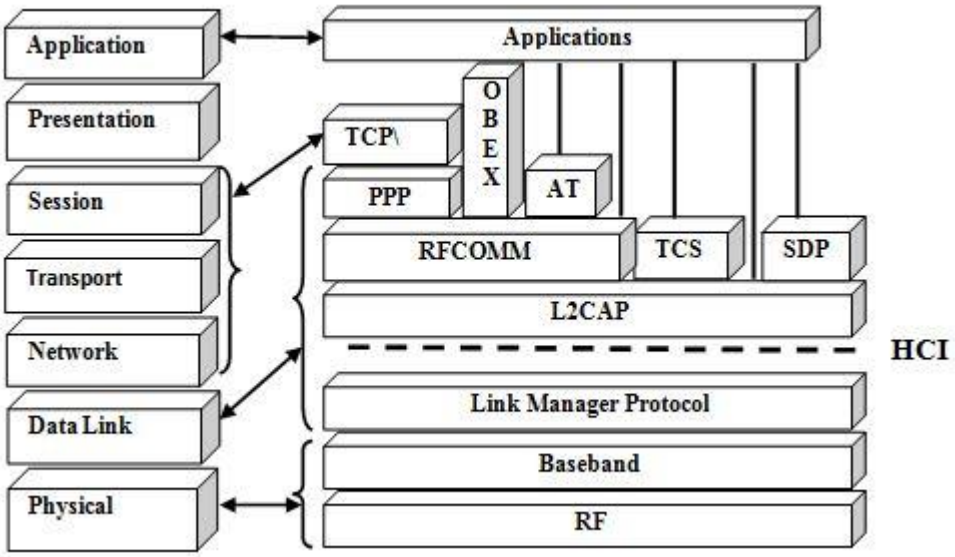


Figure 2. 2 layers of Bluetooth Protocol [12]

2.2 Some advantages of Bluetooth:

Due to its attractive features, short-range wireless transmission, small size, low cost, low power consumption, BT has become an efficient world wide solution for wireless personal area networking. Beside these features, BT includes more advantages:

1. It is used as a cable replacement: via this technology various types of cabling required to connect two or more different or similar devices such as keyboard, headsets, printers, mouse etc.

2. synchronization of Wireless: with bluetooth enabled devices it automatically synchronized for example synchronization of address book contained in laptops,cellular phones.

3. Internet connectivity: any Bluetooth-enabled device with Internet access can share Internet access with another Bluetooth device.

4. handle data and voice in the same time so that enables users to have a variety of uses such as printing ,synchronization with PC and laptops, accepting voice calls through hands-free headsets etc.

2.3 Bluetooth radio spectrum

Radio spectrum is open to the public without the need for licenses [9]. The unlicensed radio band is worldwide available. This band, the Industrial, Scientific, Medical (ISM) band, is available from 2400 MHz to 2483.5 MHz (2.4 GHz-2.4835 GHz). The universal availability and freedom of the ISM band have some unfavorable effect. This unfavorable effect is noisy channels. Noisy channels can be occured by other wireless radios operating on the same time slot and in the same frequency of the Bluetooth radio. Consequently, an amount of interference could be added by those radios, thus the Bluetooth wireless link performance will be affected. To mitigate the interference effects some options are available for the Bluetooth radio [7]. One option is interference suppression, which demands appending more components and circuitry to deal with undesired frequencies. Another option is Avoidance, which is achieved in time or frequency. Time avoidance is proper in the case of a pulsed jamming source of interference. Frequency avoidance is more feasible because the existed free bandwidth in the ISM band is about 80 MHz and most radio systems are band-limited which rises the probability of finding a part of the ISM spectrum where no prevail interference. This reason makes the Frequency Hopping (FH) is the best channel accessing solution for the Bluetooth radio. BT technology relies on the Frequency Hopping technique to preserve the confidentiality of its wireless data as Bluetooth encryption is not trusted any more.

2.3.1 Frequency hopping:

As other standards of data communication systems such as WiFi and BigZee, BT radio system works in the free unlicensed ISM band [7]. So that there is a probability of sharing the same band by other wireless radios. Bluetooth technology uses frequency-hopping spread spectrum technology to avoid interference cases. It is also

proposed to improve Bluetooth performance by avoiding noisy channels. Basically, FH is a technique by which the signal hops across several different frequency channels. This is all based on a small network called piconet ,where there is a single master device and 1-7 slave devices. You can have multilple piconets in the room and they should not interfere each other because of this ability of the signal hop. In the ISM 2.4 GHz band, Bluetooth uses 79 radio frequency channels each spaced 1 MHz apart as shown in fig. 2.3. Each packet utilizes a single 625 time slot. However, it can be spanned to cover up to five slots. the Bluetooth produces about 1600 hops per second which are selected in a pseudo-random way [7].

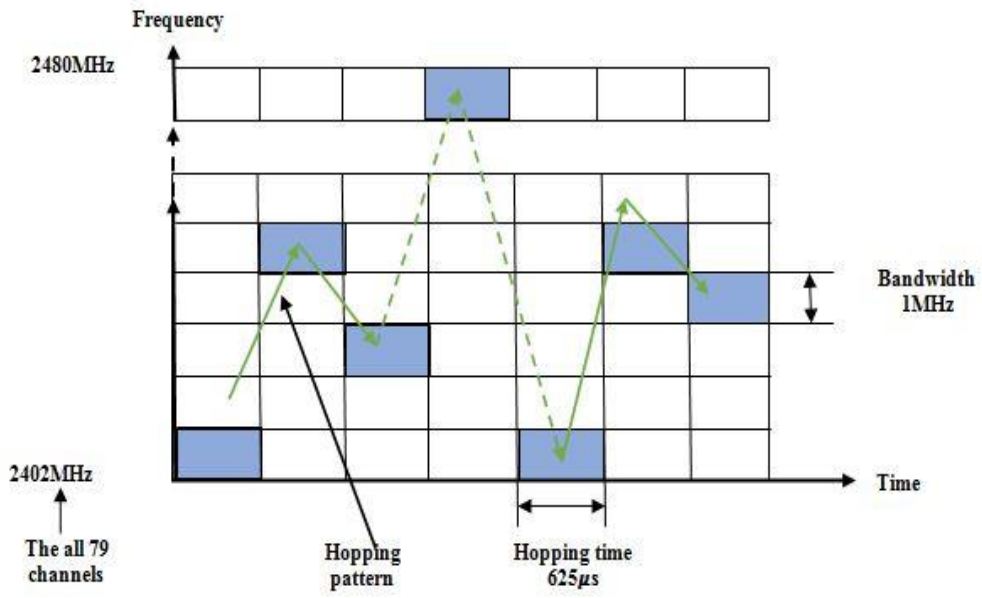


Figure 2.3 Transmitting Bluetooth wireless packets over multiple channels [7]

2.3.2 Frequency bands and channel arrangement

Here in this section some characteristics of BT technology is explained:

Within 2.4 GHz ISM band, Bluetooth defines 79 RF channels by 1 MHz spacing, and 40 RF channels with 2 MHz spacing for Low Energy transmissions [8]. The spacing starting from 2402 MHz to 2480 MHz as in fig. 2.4. This arrangment of BT channals provides a guard band. Between the guard band 79 channals are possible.

The spacing can be ruled by the following equation : f = 2402 + K MHz; k = 0, 1,…,78, and f=2402+k*2 MHz, k=0, … ,39; for Low Energy transmissions. Where f (in MHz) specifies the channel frequency that will be used for transmission. K is an index that specifies the channel number. The index K can be determined by a hardware unit called The Hop Selection kernel. In Hop Selection kernel a large number of pseudo-random hopping sequence is assigned. The Hop Selection kernel chooses a certain hopping sequence based on the MAC address of the Bluetooth device.

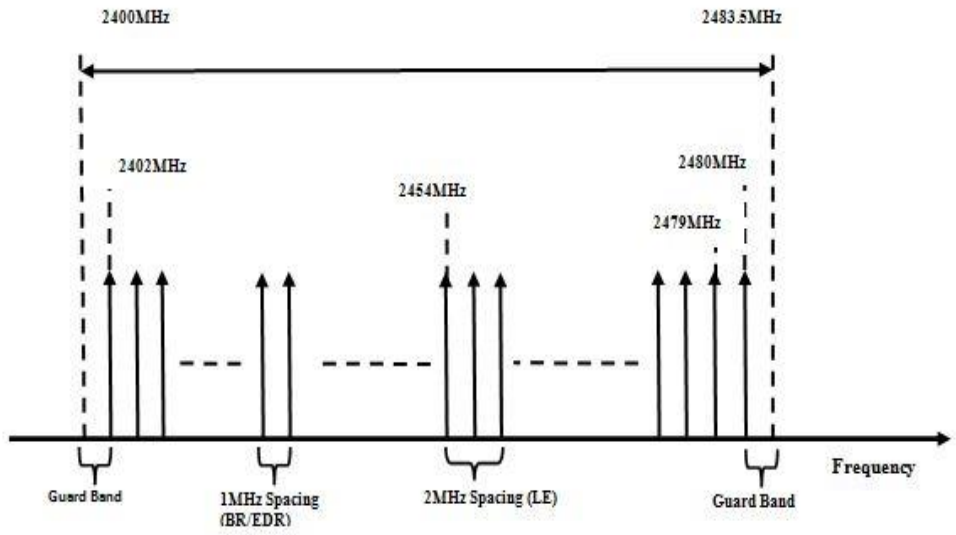


Figure 2.4 Bluetooth Frequency Bands and RF Channels [8]

625 μ□second is a nominal dwell time or time slot used for hopping channel. The frequency hops occur at a standard hop rate of 1600 hops/s [7, 8].

2.4 Bluetooth packets

Bluetooth system employs packet-based transmission as shown in fig. 2.5. The bit stream is fragmented into packets. only a single packet can be sent in each slot. All packets share the same format except the payload section. The payload section his length is variable and depends on the available user data. The packet format starts with an access code, followed by a packet header, and ends with the user payload.

72 bits	54 bits	0-2745 bits
Access code	Header	Payload

Figure 2.5 Bluetooth wireless packet format [7]

The access code is used as a direct-sequence code in certain access operations. It includes the identity of the piconet master. The packet header contains link control information. It is limited to 18 information bits in order to restrict the overhead [9]. Lastly, the payload section will include the data that is being transmitted.

The physical RF channel is divided into time slots. In which the time slots associated with a specific hop frequency. The basic time slot is 625μs (1600hops/s) in length. In a piconet, data is transmitted between the Bluetooth devices by packets that are located in their particular time slots.

A single wireless packet can be sent on each wireless channel. A new wireless channel is chosen in a pseudo random way in the next Packet. Packages sent by either party must be installed when received by the other party. This is known as Acknowl- edgement/Negative Acknowledgement (ACK/NACK)[7].

2.5 The modulation scheme

A binary modulation scheme was chosen for more robustness since in ISM band the bandwidth of the signal is limited to 1 MHz in FH systems. With this restricted bandwidth, the data rates are limited to 1 Mb/s. With this data rate the Bluetooth uses Gaussian- frequency shift keying (GFSK) modulation. For enhanced data rate (EDR), Bluetooth uses Phase Shift Keying (PSK) modulation with 1 Mb/s. Two different PSK modulation schemes are used: For 2Mb/s, $\frac{\pi}{4}$ -DQPSK modulation is used, and for 3Mb/s, 8-DPSK modulation is used [8].

2.6 Network topology

The Bluetooth devices to be connected together, Bluetooth standard utilizes a master and slave model [7]. That model allocates up to seven devices to be connect together in a small net called a piconet. In the picont there is one device called a master and the other devices called slaves, see fig. 2.6_a. The responsible for initiating a connection is the master device. All the slaves are synchronized in all aspect with the master. Apiconet has one master. A master can be called a primary station and slaves called secondary station. Extra slaves can be connected to other piconet. If a slave is shared by two masters, it will also be shared between two piconets, so that slaves can be a part of multiple piconets. All devices on a piconet follow the Master's frequency hopping sequence and timing.

2.6.1 Point to point and point to multipoint piconets

The communication between a master and slaves can be one to one or one to multypoint. In fig.2.6_a, the piconet on the right illustrates a point to point connection with only one Slave. The piconet on the left illustrates a point to multipoint connection with three Slaves talking to the Master [11]. The communication takes place between the master and slaves but no direct communication takes place between slaves. The formation of a piconet is governed by two factors, address of each BT device and clock associated with each device. Every device in a piconet has been assigned with 48 bit. Address which is similar to an ethernet address. The address field is divided into three parts in the lower address part lap is used for the purpose of piconet identification, error checking, and security checks. Every device has 28 bit clock which is called a native clock its freqency is 3200 pulses per second because this exactly twice the normal hopping rate of 1600 hops per sec. As mentioned before, a piconet can have up to seven active slaves at any instant of time. In order to identify a slave , each one is associated a locally unique active member address M address. If a BT device is not associated with any piconet , then it is said to be in standby mode.

2.6.2 Scatternets

The specification defines the number of Slaves in a piconet to seven, all Slaves only communicating with a shared Master. However, a greater number of network members (a large coverage area) can be achieved by linking piconets into a scatternet, where some devices are members of more than one piconet (see Figure 2.6_b). That is means, different piconets can be joined simultaneously into scatternets. They may even overlap each other [8]. The data rate within a single piconet is higher than data rate between scatternets. A scatternet includes all the interconnected piconets as shown in fig 2.6_b. In fig. 2.6_b, the scaternet is combining of 2 piconets as an example. A slave in the first piconet can act as a master in the second piconet. It will receive the message from the master in the first picont by acting as a slave and then delivers the massage to the slaves in the second piconet as a master. So by increasing the number of piconets, the possiblity of collisions increases. This will results in degradation of performance, therefore a device can participate in two or more piconets by means of the time sharing the process. To do so, it must use the associated master address and proper clock offset. A BT architecture device can acts as a scatternet in that the utilization of bandwidth is not optimal. This happen because a device changes its role and takes part in fifferent piconets. Another important issue is that a device would be missing when it participate in more than one piconet. If a master of one piconet temporarlily becomes a slave in some other piconet, then it will be missing from its own piconet for that much time this reduce the quality of BT link.

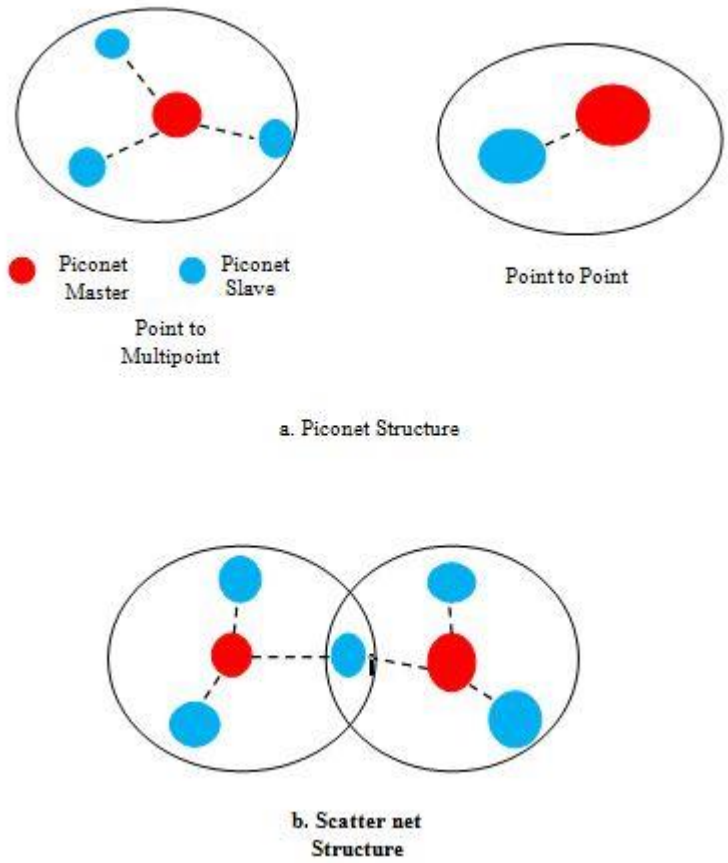


Figure 2.6 Bluetooth Network topology [11]

CHAPTER 3

HILBERT–HUANG TRANSFORM (HHT)

Conventional data-analysis methods are based on linear and stationary assumptions. Over the last few years, new methods have been introduced to analyze nonstationary and nonlinear data; Stationarity means that the statistical properties such as mean, variance, and covariance of the process do not change over time. Non-stationary data are unpredictable and can not be processed. For reliable results, the non-stationary data needs to be transformed into stationary data. Hilbert–Huang transform (HHT) helps to produce a physically meaningful representations of nonlinear and non-stationary data. It has been the most distinguished adaptive development to decompose signals. An adaptive means that the definition of the basis has to be data-dependent; in other words it is an a posteriori-defined basis. An adaptive basis is necessary condition to represent nonlinear and nonstationary data. HHT is originated from applying Empirical Mode Decomposition (EMD), a recently developed data adaptive method for nonlinear and non-stationary signal analysis, with Hilbert transform (HT) [19, 20]. Huang *et al.* Was the first to introduced the EMD and HSA as a new adaptive data analysis tool. Cocequently, the National Aeronautics and Space Administration (NASA) assigned the HSA and EMD together as the Hilbert–Huang transform (HHT). The EMD and HSA were introduced as an adaptive data analysis tool [21]. The conclusive goal of the whole effort was to have a truly time-frequency representation for nonlinear and nonstationary data or the Hilbert spectral representation. EMD was the smoothing step to set the data, then the HT could be used to define the instantaneous characterristics such as amplitude, frequency, and phase for the time-frequency representation (TFR). The time-frequency representation of any signal gives information about how the signal spectral content

changes with time, thus providing an ideal tool analyze non-stationary signals. The need for a TFR comes from the inadequacy of either time basis or frequency basis analysis to completely describe the nature of non-stationary signals [20]. There are many methods have been devised to obtain energy density as a function of time and frequency. Some of these methods are: the short time Fourier transform (STFT), the Wavelet Transform (WT), and the Wigner-Ville (WV) distribution (WVD) [20]. On the whole, some of them represent a significant defect; for instance, WVD, which is a quadratic time frequency representation, which has been done by a multiplicative comparison of a signal with itself and then expanded in different directions about each point in time, it gives better accuracy of time and frequency but inserts negative values, which leads to negative energy, thus this is not physically possible [20]. Another pupular method is STFT, the signal frequency components localized by modulating it with a window function. However, the selection of the window to have the same shape (as the analyzing signals) improves the performance a bit slightly. In WT, the main problem that it is not possible to process the phase information of the signal [20]. Among the mentioned time frequency representation methods Wavelet transform is mainly used in image processing and WVD has a vast domain of applications. The STFT is quite used to analyze audio signals in time-frequency domian when amplitude and phase information are required. For appropriate reconstruction of the signal after some additional processing in time-frequency domain is rather difficult without the phase information. That is why, the STFT is taken into account to compare with the HS as time frequency representation methods of audio signals. In this work, the performance of Hilbert spectrum (HS) is introduced as a technique for time frequency representation of BT signals. We will introduce a the HS which requires two steps in analysing the data as follows: The first step is to prepare the data by the empirical mode decomposition method, by which the data will be decomposed into a finite number of intrinsic mode function components. The second step is to construct the energy frequency time distribution by applying the Hilbert transform to the decomposed IMFs in order to design the Hilbert spectrum [22].

3.1. Empirical mode decomposition

As it is mentioned before HHT is a blend of two parts. The first part is an empirical mode decomposition which used to decompose a signal into sub-signals called intrinsic mode functions (IMF). It is viable for nonstationary and nonlinear data [19]. The EMD is a systematic way, involves approximation with splines, to extract IMFs from a signal [23] based on a sifting procedure [24]. The IMF represents a simple oscilation that have the same number of extrema and zero-crossings. IMF can have a variable amplitude and frequency as functions of time. The generated IMFs which are a series of stationary and linear sub-signals must satisfy the two following conditions [16]:

1. For each generated IMF data set, the number of zero crossings and the number of extrema (maxima or minima) must be either equal or at most different of each other by one.
2. At any point, the mean value of two envelops (upper and lower envelops) is zero; where the upper envelope defined by the local maxima and the lower envelope defined by the local minima of the generated IMF [19, 22, 25].

The maximam points of the signal are connected via cubic spline line function in order to define an envelpe which is called the upper envelope. The same process is applied for the minimum points of the signal to define the lower envelope. In this manner, The upper and lower envelopes will cover all the data between them. After getting the upper and the lower envelopes, obtain the mean of the two envelopes. The next step is taking the difference between the original data and their mean . The next section explains the procedure in details.

3.1.1 The steps of EMD

The first step of the EMD is obtaining of the local maxima and minima of the input signal. All the local maxima and all the local minima are connected by a cubic spline, in order to form the upper envelope and the lower envelope, respectively, to the original input signal *x(t).*

In the second step, determine the mean, *m (t),* of the upper and the envelopes. Define the original signal and the mean as follows

$$h_1 = x(t) - m_1(t) \tag{3.1}$$

The third step is an investigation based on the given conditions to evaluate whether the resultant $h_1(t)$ represents an IMF or not. The obtained $h_1(t)$ can be also evaluated by using a Cauchy type of convergence test. This test is using to determine a stoppage criterion. Specifically, the normalized squared difference between two successive sifting operations is required in this test, and can be defined as [26]

$$STD_k = \frac{\sum_{t=0}^{T} |h_{k-1}(t) - h_k(t)|^2}{\sum_{t=0}^{T} h_{k-1}^2(t)} < STD_0, \tag{3.2}$$

where STD_0 is predetermined value.

If $h_1(t)$ does not represent an IMF, then $h_1(t)$ is used to be the original signal; and the above steps are repeated as

$$h_{1k} = h_{1(k-1)}(t) - m_{1k}(t), \tag{3.3}$$

until $h_{1k}(t)$ represents an IMF. Once the first IMF is obtained, which is equal to $h_{1k}(t)$, a variable vector can be defined as following:

$$C_1(t) = h_{1k}(t), \tag{34}$$

where $C_1(t)$ represent the first IMF extracted from the original input data x(t). Due to this extraction, a residue $r_1(t)$ from $x(t)$ can be calculated as follows

$$r_1(t) = x(t) - C_1(t). \tag{3.5}$$

In the fourth step, the first residue $r_1(t)$ that is obtained in the third step is used as the original signal. The previous steps need to be repeated based on new original signal, $r_1(t)$, to obtain the IMFs, $C_2(t), C_3(t), \ldots, C_n(t)$. The decomposition procedure ends based on stoppage criteria. These criteria can be the satisfaction of one of the following conditions [27]:

I. The residue $r_n(t)$ which represents the last IMF $C_n(t)$ is lower than a given threshold.

II. The residue $r_n(t)$ is a monotonic function.

The original signal can be reconstracted as the summation of all the obtained IMFs adding to the residue. This is given in the following formula:

$$x(x) = \sum_{i=1}^{n} C_i(t) + r_n(t) \tag{3.6}$$

.

Each IMF represents a mono-component contribution where the derivation of instantaneous amplitude and frequency provides a physical significance. The following flowchart illustrated the EMD algorithm. The inner loop in the flowchart checks whether the extracted signal represents IMF or not. If so, an IMF is generated. If not, the extracted signal is sifted till an IMF is generated. Once the IMF is obtained, it is subtracted from the original signal, and then the subtraction result is checked in the outer loop. The outer loop checks the result (residue) whether it represents a monotonic function or not.

3.2. The Hilbert spectral analysis

It's easy to apply the Hilbert transform to each obtained IMF, where the instantaneous amplitude, instantaneous phase, and instantaneous frequency can be computed.

The Hilbert transform of the real valued u(t) generates a complex conjugate *y(t)*. The direct and inverse Hilbert transforms of *u(t)* and *y(t)*, respectively, are given by [28]

$$y(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{u(\delta)}{t-\delta} d\delta = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{u(t-\delta)}{\delta} d\delta, \tag{3.7}$$

$$u(t) = \frac{-1}{\pi} P \int_{-\infty}^{\infty} \frac{y(\delta)}{t-\delta} d\delta = \frac{-1}{\pi} P \int_{-\infty}^{\infty} \frac{y(t-\delta)}{\delta} d\delta, \tag{3.8}$$

where *P* indicates the principal value of the singular integral. The integral in Eq.(3.7) and Eq.(3.8) are improper. In order to solve these integrals it is needed to evaluate them as Cauchy Principle Value (*CPV*) Integrals. The singularity must be considered and the limiting process must be applied. Having taken the Hilbert transform of the function *u(t)* , an analytic signal can be define as

$$z(t) = u(t) + iy(t). \tag{3.9}$$

3.2.1 Signal characteristics

After having implemented the EMD and generated all the IMFs of the input signal, the Hilbert transform is needed to be applied to each extracted IMF to generate its instantaneous characteristics. These instantaneous characteristics are expressed as the instantaneous amplitude, instantaneous phase, and Instantaneous frequency [14]. Applying the Hilbert transform to the IMFs yields the following analytic signals:

$$Z_m(t) = i_m(t) + jq_m(t), \tag{3.10}$$

where $Z_m(t)$ represents the analytic signal corresponding to the $m_{th}IMF$, $q_m(t)$ is the $H(i_m)$. the instantaneous amplitude and the instantaneous phase corresponding to the $m_{th}IMF$ can be computed, respectively as following:

$$a_m(t) = [i_m^2(t) + q_m^2(t)]^{\frac{1}{2}}, \tag{3.11}$$

$$\theta_m(t) = tan^{-1}\left[\frac{q_{m(t)}}{i_m(t)}\right], \tag{3.12}$$

where i(t) and q(t) represent the in-phase and quadrature components of the complex signal, respectively, [14]. Based on these characteristics, the analytic signal that is given by Eq. (3.10) can be expressed as

$$Z_m(t) = a_m(t)e^{i\theta_m(t)}. \tag{3.13}$$

The Instantaneous frequency (IF) of the $m_{th}IMF$ can be computed as the derivative of the corresponding phase to that IMF as follows

$$f_m(t) = \frac{d\vartheta_m(t)}{dt}, \tag{3.14}$$

where $\vartheta_m(t)$ is the unwrapping of the instantaneous phase $\theta_m(t)$. The derivative in Eq. (3.14) is considered as a discrete instant of time t. This derivative generates abrupt fluctuations of the IF so that a nonlinear smoothing is needed [29]. Some types of smoothing filters are used such as moving average filter [29], Savitzky-Golay filter [30]. The concept of IF does not have a physical meaning, if it is applied to arbitrary signal. This signal must be decomposed into a series of mono-component contributions, and then the IF will have a physically meaningful if it applied to these mono-component contributions. It can be observed that the corresponding IFs to each IMF are different. Therefore, EMD method is considered as an effective decomposition technique of non-linear and non-stationary signals based on their frequency characteristics. This property identifies each frequency component of the analyzed signal, and localizes it in both time and frequency scales. That yields spectra at each sampling point [29].

3.3. Hilbert spectrum

Having obtained the characteristics of each IMF, the Hilbert spectrum (HS) can be generated. The HS represents a three dimensional plot illustrates the energy distribution as a function of time and frequency [20, 29]. It is defined as the three dimensional plot of time-frequency distribution of amplitude [31]. It is also constructed as Hilbert amplitude spectrum [20].

The instantaneous frequency (IF) of each IMF is obtained from the derivative of the corresponding unwrapped instantaneous phase. This derivative generally generates abrupt fluctuations of obtained instantaneous frequency so that smoothing techniques such as the moving average smoothing filtering [20, 32] and the median smoothing [27] can be used to get rid of such fluctuations. Here the sgolayfilt function is applied.

The reason behind the usage instantaneous frequency and instantaneous amplitude to generate the HS is that the non-linearity and non-stationary can be handled better than using the traditional paradigm of constant frequency and amplitude [33]. Having performed the Hilbert transform on each IMF, the time series original data can be expressed as [26, 34]:

$$x(t) = \Re\left\{\sum_{i=1}^{n} a_i(t)\, e^{j\int w_i(t)dt}\right\} \tag{3.15}$$

where R (.) represents the real part of the resultant analytic signal and n is the number of generated IMFs. The instantaneous amplitude and the IF can be represented as a function of time based on Eq. (3.15). The following steps represent the procedure to construct the HS [26, 30, 34]:

Determine the maximum and the minimum values of the whole instantaneous frequencies IFs as $IF_{max} = Max(IF_1, IF_2, \dots, IF_m, \dots, IF_n)$ and $IF_{min} = Min(IF_1, IF_2, \dots, IF_m, \dots, IF_n)$ calculated from all instantaneous frequencies [24, 32]. Generate scaling factor $\eta = nf/(IF_{max} - IF_{min})$, where the IF is normalized between 0 and a certain positive real number nf. Each IF vector is multiplied by the scaling factor η. Determine the bin spacing of the HS as $bs = nf/B$, where B is the number of desired frequency bins. Calculating the HS of the signal is constructed from the superposition of the individual HSs of all IMFs as :

$$H(l,t) = \sum_{i=1}^{n} H^{(i)}(l,t) \tag{3.16}$$

where $H^{(i)}(l,t)$ is the HS of the i_{th} IMF, *l=(1,2,…, B).* Each entry *H(l,t)* in the HS is obtained by the weighted sum of the instantaneous amplitudes of all IMFs at l_{th} frequency bin based on the following formula:

$$H(l,t) = \sum_{i=1}^{n} a_i(t)\, w_k^{(l)}(t) \tag{3.17}$$

where the weighted factor $w_k^{(l)}(t)$ takes 1 if $\eta \times IF_m(t)$ falls within l_{th} band, otherwise is 0. Having computed all the HS elements over the frequency bins, the resulted space is time frequency representation of the instantaneous signal spectrum. The phase information corresponding to each cell in the HS space can be obtained by the same manner that is used to construct the amplitude. The following formula illustrates the instantaneous phase of all IMFs at l_{th} frequency bin [32]:

$$\varphi(l,t) = \sum_{i=1}^{n} \theta_i(t)\, w_k^{(l)}(t) \tag{3.18}$$

There are some considerations when the HS is constructed. If the energy density is desired to be represented by squared amplitude, then the values of the squared amplitude is used to construct the Hilbert energy spectrum. If just the visualization is desired, then the HS can show where most of the signal energy is concentrated in the time frequency plane [35].

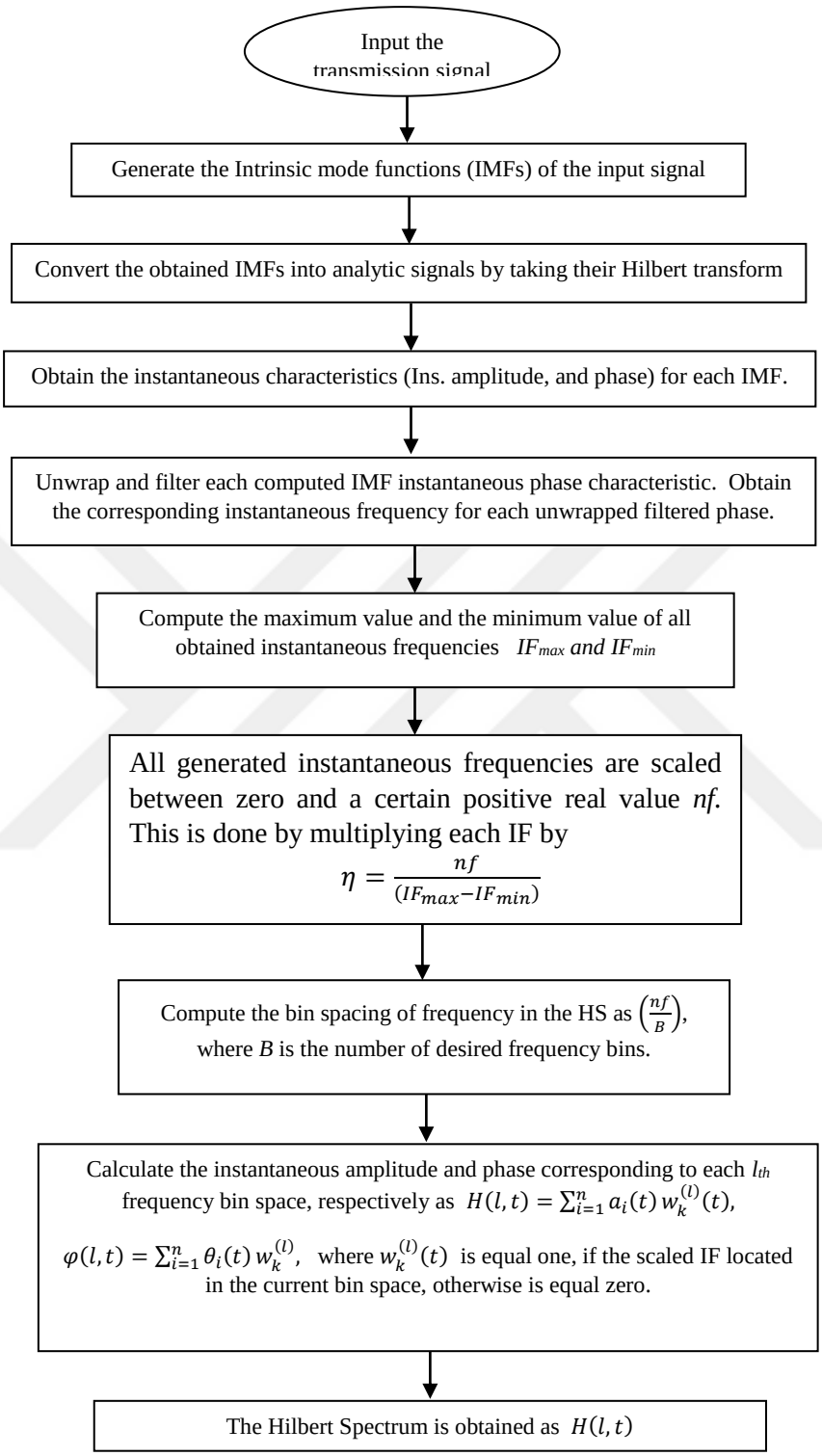


Figure 3.1 A flowchart of generating Hilbert Spectrum

CHAPTER 4

TRANSIENT SIGNAL DETECTION

In RF fingerprinting techniqe, detecting the transient signal is considered to be the most important stage. Since the transient signal contains the unique features that ditinguish the cell phone devices from each other, so precise detecting is needed for high accuracy of classification and identification devices.

The transient signal occures when the devcice is turned on. It occures in a very short duration of time. This duration about 13 μs in Bluetooth (BT) signals. The transient signal has been represented by the attributes to the variety of transmission signal sources. Some of these sources are modulator subsystem, phase-lock-loop systems, radio frequency power amplifier, modulator subsystem, and antenna [16].

In the previous studies, the transient signal was detected by means of many approaches. The start of the transient of a transmission signals can be detected based on the signal features. These features are extracted from the time domain or frequency domain, where the time domain is expressed as the signal amplitude, the frequency domain is represented by the signal frequency or phase [14]. According to [14] some methods such as the Threshold, Bayesian Step Change Detector, and phase based method were used to detect the start point. Threshold and Bayesian Step Change Detector the amplitude characteristic was exploited to detect the start point; whereas phase based method exploited the phase characteristic to detect the start point. In [18], the transient endpoint of the transmitted signal was detected by the characteristic of the transient energy trajectory from Higher order cumulants. In [16], the endpoint was detected in an experimental manner. In [13] the start point was detected by calculating the variance of the signal phase.In [36] the transient detection achieved by using the Gabor representation.

In our work, we detected the start and end points by means of energy envelope method. By using this technique there is no difficulty to select appropriate threshold value for all BT signals. The sellected thresholds are assigned to detect the start and the end points for all cosidered signals without tunning. That is helpful when huge data is considered.

4.1 Start transient detection

The start of the transient of a transmission signal can be detected based on the signal features. These features are extracted from both the time domain (amplitude) and the frequency domain (frequency and phsae) [14]. The start point of the transient signal can be detected via many techniques. We will introduce some of them to detect the start of transient signal.

4.1.1 Threshold detection approach

The threshold is defined based on an accumulating confidence decision algorithim to achieve reliable detection [37]. In the threshold detection method, the variance of the signal amplitude represents the feature to be extracted. This feature is obtained as follows [13]

$$D(t) = E + 1 - H \tag{4.1}$$

where E represents the Euclidean dimension; is given as zero for point, one for line, and two for space [38]. In our application E is assigned to be one.

D(t) Represents the variance dimension or fractal dimension. *D(t)* is an important characteristic since it includes information about their geometrical structure at various scales. *D(t)* is fall between highly correlated signal portion at *D(t)=1*, and uncorrelated noise at *D(t)=2*. (*H*) represents the Hurst exponent. (*H*) is a statistical measure used to classify time series. (*H*) is related to the variance dimension or fractal dimension such that as H increases from *0 to 1*, *D(t)* decreases from *1 to 2*. [38].

On the other hand, The Hurst exponent (*H*) can be from the signal as following:

$$H = \lim_{\Delta t \to 0} \frac{1}{2} \frac{\log[var(\Delta X_{\Delta t})]}{\log(\Delta t)} \tag{4.2}$$

where *Δt* represents the time increment, and $\Delta X_{\Delta t}$ is the difference between two samples of the discrete time series signal, *Xi : i=1,2,3,…,n.,* that are separated by *Δt.* H can estimated by different methods such as

$$var(\Delta X_{\Delta t}) \approx |\Delta t|^{2H} \tag{4.3}$$

The Hurst exponent can be also calculated by the following procedure [39, 40]:

Calculate mean value *m* of the discrete time series signal *Xi* as

$$m = \frac{1}{n} \sum_{i=1}^{n} X_i \ . \tag{4.4}$$

Calculate mean adjusted series *Y*

$$Y_i = X_i - m. \tag{4.5}$$

Calculate cumulative deviate series *Z*

$$Z_q = \sum_{i=1}^{q} Y_i \,, \qquad q = 1, 2, \dots, n \tag{4.6}$$

Calculate range series *R*

$$R_t = max(Z_1, Z_2, \dots, Z_q) - min(Z_1, Z_2, \dots, Z_q). \tag{4.7}$$

Calculate standard deviation series *S*

$$S_q = \sqrt{\frac{1}{q} \sum_{i=1}^{q} (X_i - \mu)^2} \ , \tag{4.8}$$

Where μ represents the mean value of the series range *Xi to Xq*

Calculate rescaled range series (*R/S*)

$$\left({}^{R}/_{S}\right)_{q} = {}^{R_q}/_{S_q} \quad q = 1, 2, \dots, n. \tag{4.9}$$

Hurst found that (*R/S*) scales can be expressed by power-law at any point as

$$\left({}^{R}/_{S}\right)_{q} = q^{H}, \tag{4.10}$$

so *H* can be represented as the following logarithmic relationship

$$H = \frac{\log\left({}^{R}/_{S}\right)_{q}}{\log q}. \tag{4.11}$$

Having obtained the fractal dimension based on the Euclidean dimension and the Hurst exponent of the given transmission signal, the start of the transient portion can be detected according to the following inequality:

$$|D(t) - \mu| > (\tau \times \mu) + \sigma, \tag{4.12}$$

where τ is an experimentally established threshold, μ is the noise portion mean, σ represents the noise standard deviation.

4. 1.2 Bayesian change point detection

The Bayesian approach is a method by which the start of the transient portion of a transmission signal can be detected. This approach is contingent upon the learning process, by which the previous data are updated to generate new data [14]. The Higuchi's method is used, in the Bayesian detection technique, to calculate the fractal dimension [13, 41]. The fractal dimension can be determined in a reliable way, if this method is applied appropriately. The importance of the fractal dimension comes from the characterization of the degree of the data correlation [41]. The Higuchi's method can be applied to an uniform time series, *x(i), i=1, 2, …, N* .
A new series can be generated from the given original uniform time series as:

$$x_k\hat{}m: x(m), "\ " x(m+k), "\ " x(m+2k), \ldots, "\ \ " x(+[(N-m)/k]k), \qquad (4.13)$$

Where $m = 1,\ 2,\ 3, \ldots,\ k$; and k is integer number, and [ ] denotes the integer part. The length of the generated series is defined as:

$$L_m(k) = \frac{1}{k}\left\{\left(\sum_{i=1}^{\left[\frac{N-m}{k}\right]} |x(m+ik) - x(m+(i-\ 1)\ k)|\right)\frac{N-1}{\left[\frac{N-m}{k}\right]k}\right\}, \qquad (4.14)$$

Take the average of all $L_m(k)$ to obtain the length of the series (*L (k) for x (i)*) for the given *k* value. If the obtained length *L (k)* behaves as a power law i.e. $L(k) \propto k^{-D}$, then the exponent D represents the fractal dimension of the given uniform time series.

The exponent *D* which is represents the fractal dimension can be obtained by taking the natural logarithm of the vectors *L(k), and k* , followed by the application of the least square method. A straight line expresses the relation between these two vectors, and its slope represents the desired fractal dimension [14, 41].

Once the fractal dimension *D* is obtained, the posteriori probability density function can be used to detect the transient portion start point as follows:

$$P\{(m)|D\} \propto \frac{\frac{1}{\sqrt{m(N-m)}}}{\left[\sum_{i=1}^{N} D_i^2 - \frac{1}{m}\left(\sum_{i=1}^{m} D_i\right)^2 - \left(\frac{1}{N-m}\right)\left(\sum_{i=m+1}^{N} D_i\right)^2\right]^{\left(\frac{N-2}{2}\right)}}, \qquad (4.15)$$

where *m=1, 2, 3, ..., N* represents the transient portion start point, *N* is the length of the fractal dimension.

For each m value, calculate the variances of the fractal dimension for the two sequences *[1, 2, 3, ..., m] and [m+1, m+2, m+3, ..., N]* . The probability is proportional to the differences between each two variances of successive sequences, Based on the fact the fractal dimension values of the noise portion is higher than that of the transient portion so, one can expect that the transient portion start point. This

point should be located at the point where we have a highest difference of successive sequences' variances.

4. 1.3 Transient detection based on phase characteristic

The phase characteristic can represent the feature of the signal. This characteristic can be exploited to examine and analyze the characteristics of radio transmitter fingerprints [42]. The usage of the phase characteristic in the transient detections and features extraction has some advantages over the usage of the amplitude characteristics. Firstly, the signal phase is not affected by the noise and interference as much as the amplitude characteristics. Secondly, in the transient stage the phase has a linear slope. This linearity makes the detection algorithm more effective [13]. In the approaches that use the amplitude characteristic such as the Threshold detection, and Bayesian step change detector an abrupt change at the transient start is assumed. This does not work appropriately where the noise transient transition occurs gradually [13]. In order to overcome this deficiency the transient detection based on phase characteristic can be used. That is because it depends on the slope of the phase instead of the signal amplitude.

The Fractal Trajectory can be obtained by taking the Hilbert transform to the transmission signal in order to convert it into an analytic signal. The generated analytic signal represents a complex signal whereas the transmission signal is a real valued signal. Exploiting the complex values of the analytic signal, its instantaneous phase can be calculated as the arctangent of each complex entry. The signal instantaneous phase that is calculated from the analytic signal is unwrapped to delete the discontinuities [13].

In order to simplify the detection algorithm, the absolute value is taken for each entry in the unwrapped phase vector (*AV*) [13]. The variances of the unwrapped phase vector (*AV*), at the noise portion and the transient portion, are different. Along the unwrapped phase vector (*AV*), the variances can be obtained based on non-overlapping windows of size s to generate a temporary variances vector (*TV*) as following:

$$TV(i) = var\big(AV(d+1),\ AV(d+2),\ \ldots,\ AV(g)\big), \tag{4.16}$$

where $\mathrm{i} = 1,2,\ldots,{}^{\mathrm{N}}/_{\mathrm{S}}$, $g = \mathrm{i} \times s$, $d = g - s$, and *Var* represents the phase variance.

The fractal trajectory (*FT)* is obtained by taking the difference between each successive phase variances entries in *TV*.

The slope of the phase is linear only at the transient stage. Based on this, the values of the variances on the *AV* vector remain same at this portion. That means the values of the fractal trajectory (*FT*) at this portion are zero. The value of fractal trajectory can be traced based on predetermined threshold (*T*). The fractal trajectory values are evaluated based on overlapping windows with a certain length (L_{win}). The values in each window are compared to the predetermined threshold as follow:

$$FT(i),\ FT(i+1), \ldots, FT(i+L_{win}) \leq T, \tag{4.17}$$

where i=1, 2, length (FT -L_{win}*).*

Another criterion can be used to give more accuracy for the detection process. This technique is accomplished by considering successive portions of sum of signal amplitudes (*SV*). The absolute values of this difference between the value of the portion (*i*) and the value of portion (*i-1*) is compared with a predetermined threshold (τ) multiplied by the *i_th* FT value. The start of the transient portion is denoted by the index *i-1* [13].

$$|SV(i) - SV(i-1)| \leq \tau \times FT(i). \tag{4.18}$$

4.2 Detection of transient start and end points based on improved methods

In the previous sections we introduced some methods used to detect the start point of the signal transient. In this section, we will introduce two improved methods used to detect both the start and the end points of the signal transient [43]. These improved methods are Improved Bayesian change point detection and Improved Phase Characteristic.

4.2.1 Improved Bayesian change point detection

The Improvement of this approach is considered to simplify the detection process. This can be done by calculating the summations of fractal trajectory windows entries. By the summations we can get rid of calculating the posteriori probability density function [43].

In this Improved technique, the transmission signal x(t) is divided into portions based on non-overlapping windowing technique. The fractal dimension of each window is obtained where the number of windows is given by:

$$No_{win} = \left\lceil \frac{length(x)}{L_{win}} \right\rceil \tag{4.19}$$

Computing the fractal dimension by using the Higuchi's method followed by obtaining the fractal trajectory as given by Eq. (4.20), the resulted trajectory is divided into non-overlapping windows[43].

$$FT = [fd(1), fd(2), \dots, fd(No_{win})] \tag{4.20}$$

The summation of each window's entries, Eq. (4.21), is taken instead of calculating the posteriori probability density function. This is done for reducing the computation cost, simplifying the detection procedure, and distinguishing the transient portion that can be visually appears in Fig. (4-3 d). The summation widows vector (*SWV*) of the fractal trajectory is illustrated as:

$$SWV = [sum(1), sum(2), \dots, sum(No_{winFT})], \tag{4.21}$$

where *sum (i)* is the summation of i_{th} window, and No_{winFT} is the number of windows of the generated FT. The resulted *SWV* is smoothed, so that will generate a

series of data in which the transient portion is distinguished as a linear line. The difference of each two successive points in the series is taken as:

$$Diff_{AWV} = \left[\left(SWV(2) - SWV(1), \ldots, SWV(end) - SWV(end-1)\right)\right] \quad (4.22)$$

The noise and steady portions have oscillations that make the differences contain positive and negative values in these portions, whereas the differences in the transient portion contain only positive values [43].

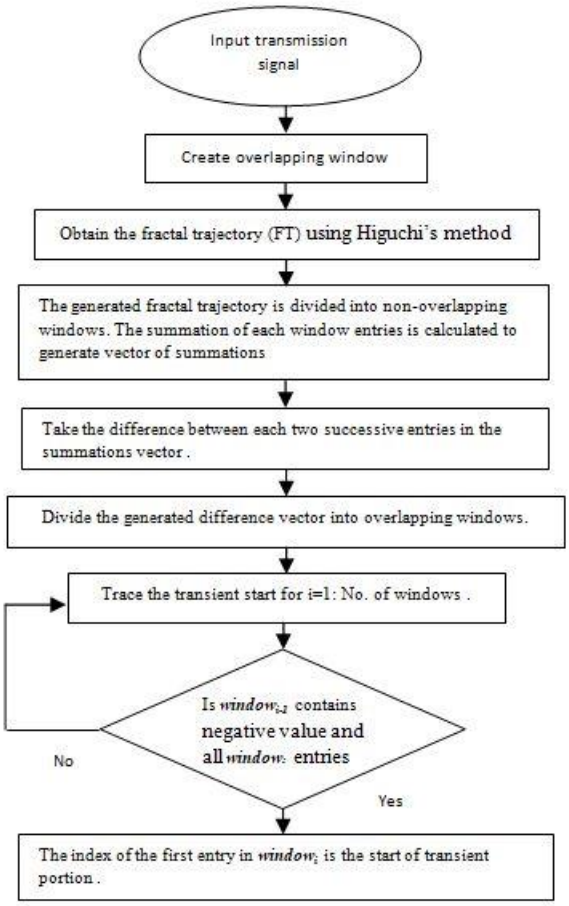


Figure 4.1 flowchart of the improved Bayesian chainge point detection

4.2.2 Improved phase based method

In this technique, after obtaining the unwrapped phase vector (AV) as illustrated in section C, the generated AV is divided into non-overlapping portions[43]. The average of each window is obtained, as given in Eq. (4.23) to distinguish the transient portion as shown in Fig . $TV = [Avg(1), Avg(2), \dots, Avg(No_{portion})]$ (4.23)

Now to detect the signal transient which has a linear slope, the difference between each two successive points of the *TV* is taken twice to differentiate the transient portion. This will generate a differences vector in which there are positive and negative values in the noise and steady states, whereas there are no negative values in the transient portion due to the linearity. The differences vector is divided into overlapping windows to detect the transient state which is represented by the first entry of the first no negative valued window [43].

In this Improved technique we avoid to use the threshold which is generally different in each signal. Moreover the average is taken instead of the variance complicated technique that is given by Eq. (4.16). This leads to simplify the procedure of the detection. Another advantage is that the signal transient is conspicuous as in fig 4-3 f. The following flowchart illustrates the steps of the detection of the transient potion first point:

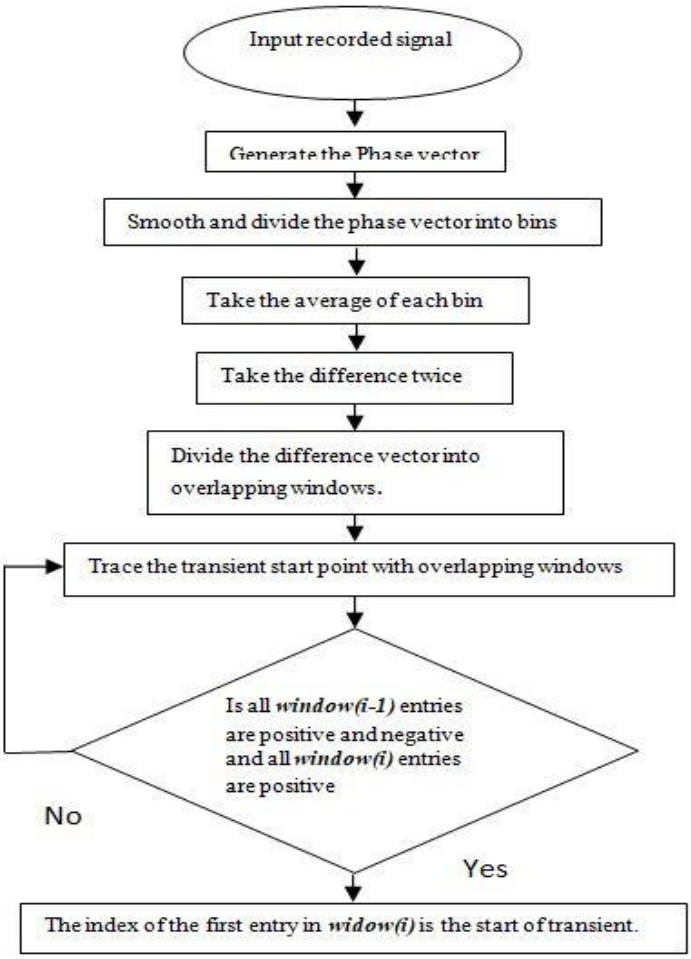
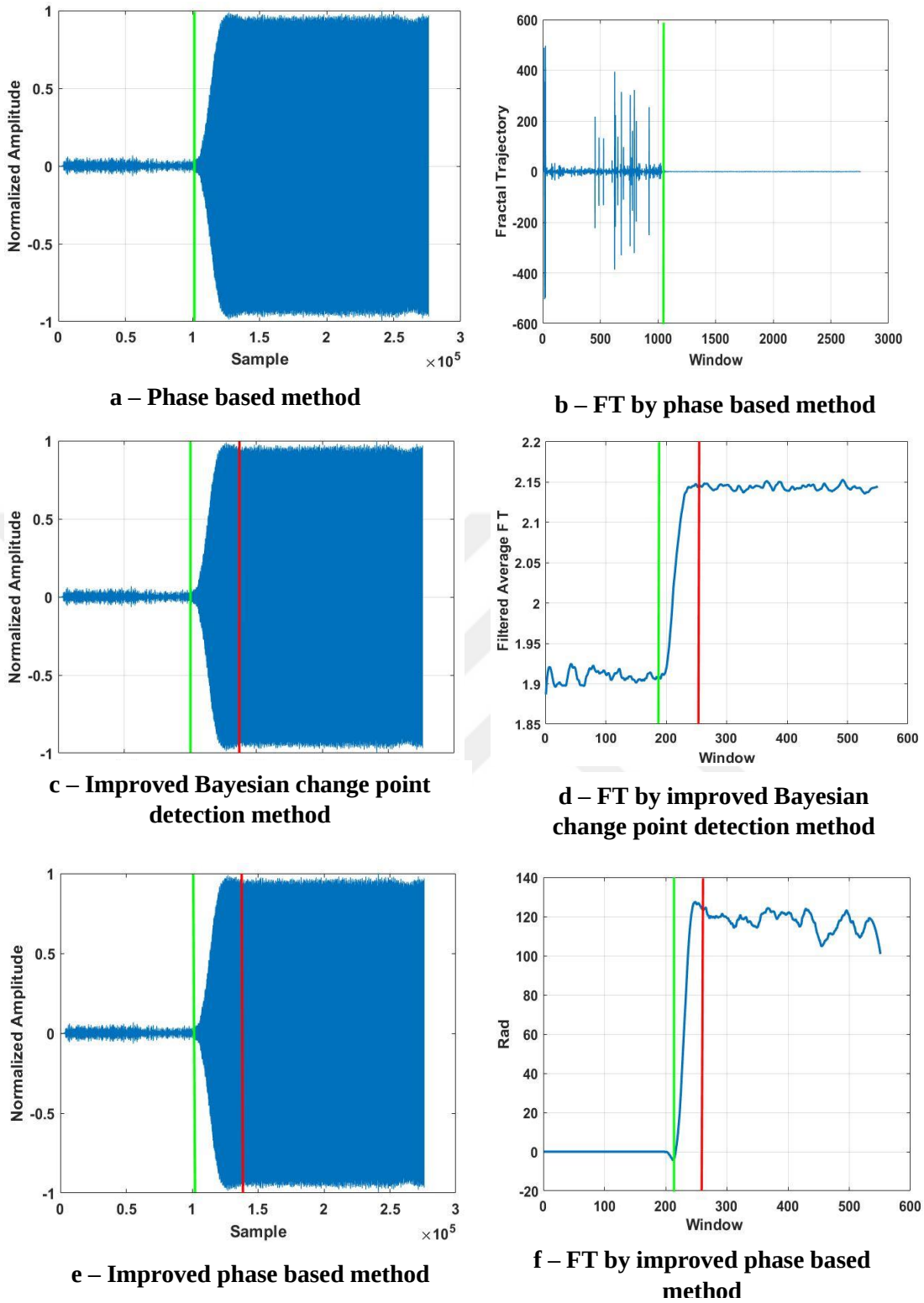


Figure 4.2 flowchart of the improved phase based method



a – Phase based method

b – FT by phase based method

c – Improved Bayesian change point detection method

d – FT by improved Bayesian change point detection method

e – Improved phase based method

f – FT by improved phase based method

Figure 4.3 comparison between the three different detection techniques

CHAPTER 5

IMPLEMENTATION

5.1. Data aquisition

Precise analysis and investigations of classification and identification procedures are the main goal of our work. This can be accomplished by using guaranteed real data and accurate procedures. Bluetooth transmission signals are captured from different cell phones brands models and different serial numbers. Fig. 5.1 shows the signal acquisition system which is designed to capture bluetooth (BT) signals from mobile phones. BT signals were captured in laboratory environment. The laboratory is isolated in an underground floor (-2) where no other devices or equipments
are switched on. The ambient temperature and humidity are very stable in the laboratory as the change in environmental conditions may affect the performance of classifications. In our work, twenty devices each with one hundred fifty records are considered. The captured signals are listed as classes in Table 5.1.

The data are captured in the laboratory and saved as text data. BT signal has a carrier frequency between 2400 MHz and 2483.5MHz. BT signals are sent to a typical modem antenna and then captured by an (TDS7404 DSO) Oscilloscope (4GHz). The distance between antenna and cell phone is about 30 cm distance in order to keep the Signal to Noise Ratio (SNR) in a high level as it is illustrated in Fig. 5.1.

Captured signals are digitized by ADC (analog to digital converter) that has at least 4.8 GHz sampling frequency. According to Nyquist theory, for a 2.4 GHz signal, ADC must be at least 4.8 GHz sampling frequency. By using this information BT signals were recorded by using 20 GHz sampling frequency. As an example, typical forms of four different collected signals are shown in Fig. 5.2. They are collected from Ipone 6S_plus, Samsung S3, Sony C4, and Huawei G5 respectivily.

Table 5.1. The Classes

Class No	Brand	Model
1	Huawei	G5 A*
2	Huawei	G5 B*
3	I-Phone	5S A
4	I-Phone	5S B
5	I-Phone	6S A
6	I-Phone	6S B
7	I-Phone	6S Plus A
8	I-Phone	6S Plus B
9	I-Phone	7 A
10	I-Phone	7 B
11	LG	G4 A
12	LG	G4 B
13	SamSung	Note3 A
14	SamSung	Note3 B
15	SamSung	S3 A
16	SamSung	S3 B
17	SamSung	S4 A
18	SamSung	S4 B
19	Sony	C4 A
20	Sony	C4 B

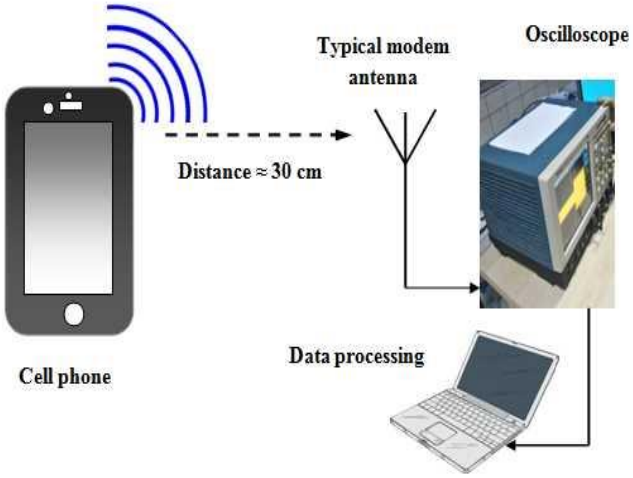


Figure 5.1. Illustration of signal aquisition system.

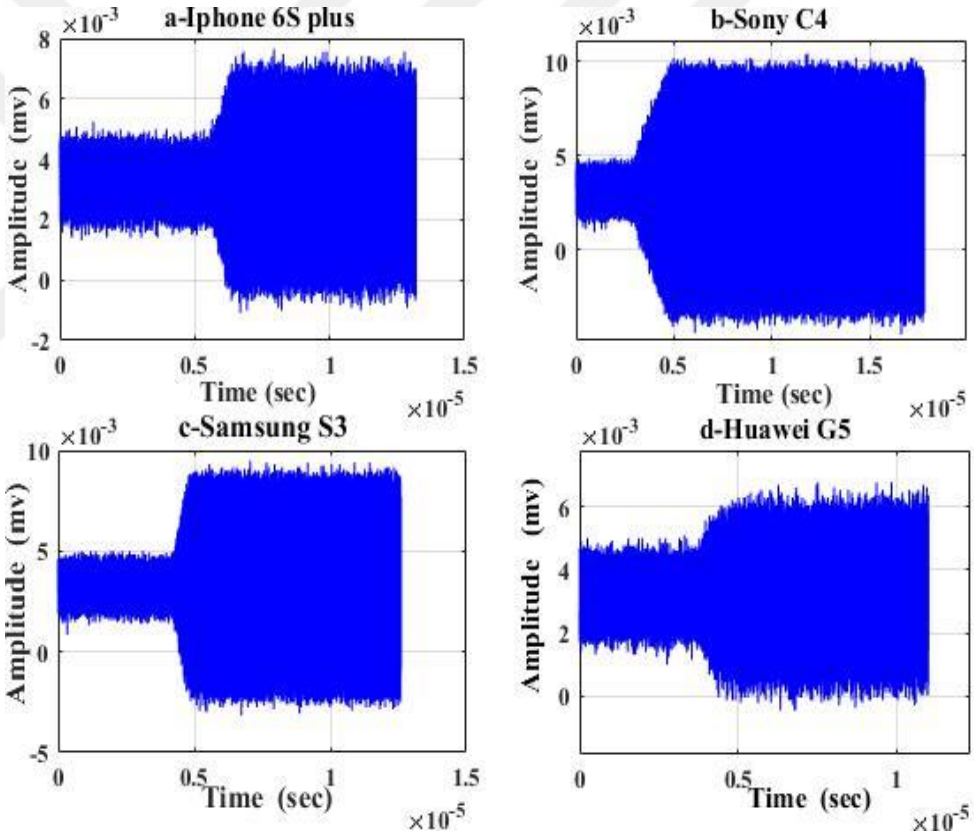


Figure 5.2 A typical form of the collected BT signal

After signal recording process, filtering is needed because it is noteced that the recorded signals have undesired frequency components [44]. These undesired frequency components are generated by oscilloscope during the recorded signal

process. To get rid of these components Band Pass Filter (BPF) is utilized. BPF which has a cut off frequencies 2400 MHz and 2485 MHz passes only ISM2400 band. Fig. 5.3 shows the same four signals that are shown in fig. 2 but after filtering process.

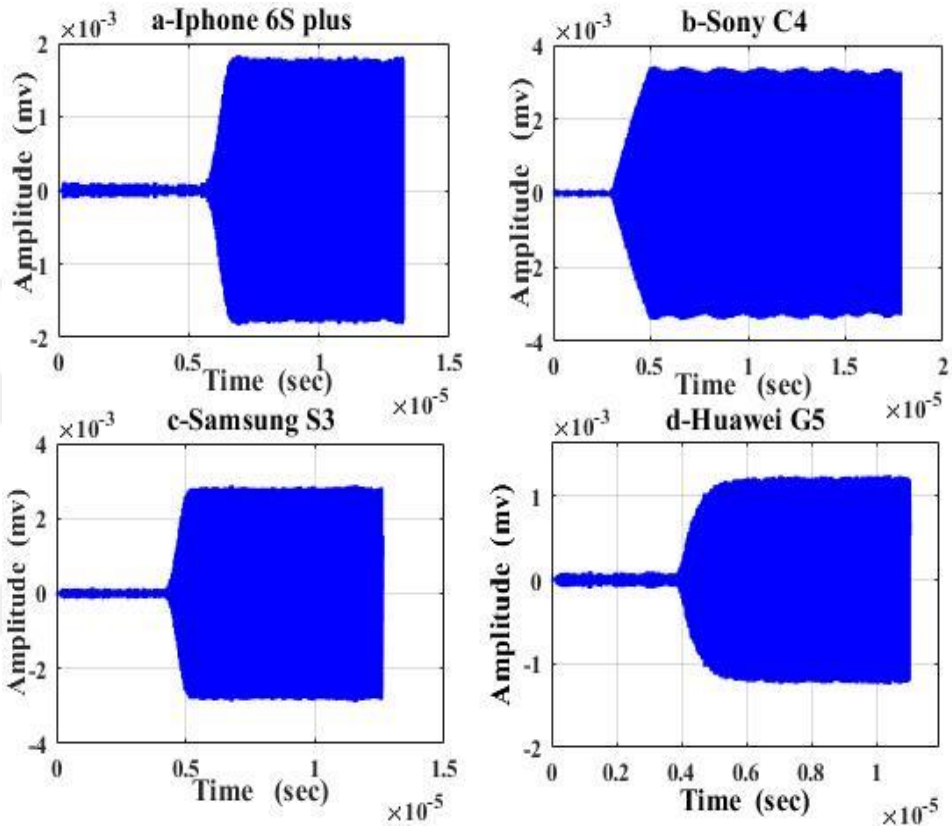


Figure 5.3 Filtered BT signals

After filtering process, signals supposed to be normalized. We need normalization because it scales all numeric variables in the range between zero and one. In this sence normalization enables us to compare data from different places more easily. Since the transmitted signals have different amplitudes, normlization is needed to shift the amplitudes of the signals in the same level. A typical form of filtered normalized BT signals of the same four different cell phoned is shown in Fig. 5.4.

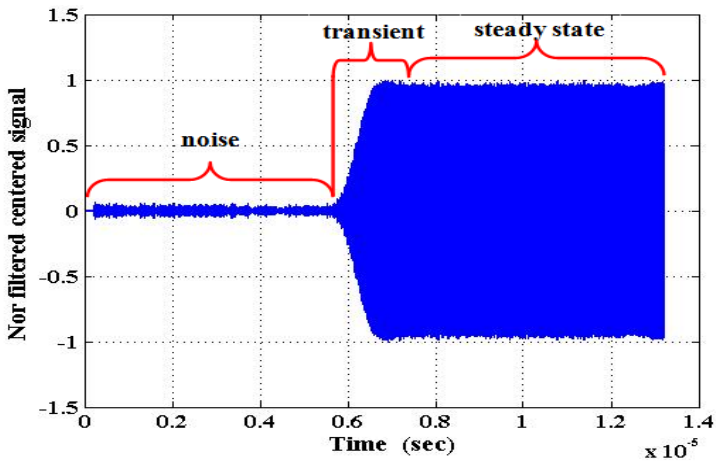


Figure 5.4 A typical form of filtered normalized BT signal

In fig. 5.4, the filtered BT signal is composed of three stages, defined as, noise stage, transient stage, and steady state stage. The gathered data are processed by means of different successive procedures in order to analyse their features. The most important features of the transmitted signals are stored in the signal transient. So that the first process of RF fingerprinting procedure is the detection of signal transient of each collected signal record. In RF fingerprinting processing, transient signal detection is an important part to extract those features and then identify specific emitter. As dicussed in chapter 4, there are many procedures to detect the signal transient. In our work the transients are detected by means of the energy envelope of the signals. Energy envelope procedure is illustrated in chapter 4. Fig 5.5 shows the dedected four different signals.

5.2 Detection of transient start and end points based on energy envelop

Since the transient extraction plays an important part in the RF fingerprinting process. So that the transient should be extracted properly in order to create precise transmitter RF fingerprint. Use of energy envelope of signals is very common technique for the detection of transients. However, it is necessary to improve the proposed techniques for both more practical implementations and lower computational loads. This requires somehow intuitive approaches to be implemented with extensive dataset. The transient is detected from the trasmitted signal via the following steps based on energy envelope:

1. The signal should be filtered and normalized.

2. Obtaining the signal envelope by calculating the instantanuous amplitude of the transmitted signal using Hilbert Transform.

3. Window the energy envelope into non-overlapping windows.

4. Take the mean of the sliding windows (to detect the start point), let the samples of the sliding windows start from i point.

5. Apply a threshold on the energy envelope and then compare the mean of the sliding window with the value of energy *E(i)* based on a certain threshold.

6. By the same technique the max local point at the end of the transient is detected.

7. The point is represented by the first positive slope of sliding window after the local maxima. The pseudo code represents some parts of the algorithm:

```
Input:    Normalized denoised BT signal.
Output:   The transient portion & the its features
Initialization: Thresholds. T1 through T6 (Assigned to fit input signals)
Steps
1:   Compute the energy envelope (EE) of the input signal
2:   Starting detection by using energy envelope
3:   for i=1: length (EE) do
4:   if  T1*mean (EE (I: i+T2)) < EE(i+T3)
5:   then Start point=i
6:   end if
7:   end for
8:   End detection by using energy envelope
9:   for  j= length ( EE) :-1: Start point  do
10:  if  mean ( EE( length( EE ):-1:j-T4) )-EE(j)>T5
11:  assign point  j.
12:  end if
13:  end for
14:  Max_point= max( EE( j:T6 ) )
15:  Windowing  EE( Max_point:length( EE ) )
16:  for k=1:# windows do
17:  if  mean( diff( window k-1) ) <0 && If  mean( diff(window k) ) >0
18:  then  the end point=index of window k(1)
19:  end if
20:  end for
```

The detection of end point of the transient signal is obtained based on the local energy envelope maximum value. This is given in line 14 in the transient detection

pseudo code. The different locations of the local maximum values of the same phone records leads to inappropriate end detection and more interference between the transient durations of different classes. In order to solve this deficiency, we exploit the transient features to re-detect the transient end. The exploited transient features are obtained close to the transient start point. These features are the standard deviation of the transient instantaneous phase, the entropy of the transient instantaneous phase, the variance of the energy envelope, and the summation of the energy envelope. We used part of the records to obtain the median of their transient durations and the mentioned features. After applying a condition of re-detection we generalized the re-detection for all records of all devices. To re-detect the transient end, following steps are needed for each phone:

1. Obtain the transient durations of some records (about the third of the records).
2. Calculate the medians of transient durations.
3. Generate the aforementioned features of each record and determine the maximum and the minimum values of each feature.
4. By means of the median and the corresponding features maximum and the minimum values, the transient end can be re-detected, for all records, by

$$E_{i,j} = S_{i,j} + \hat{\Delta}_i + R_{i,j} \tag{5.1}$$

where i represents the index number of class , j represents the index number of record, S represents the start point of the signal transient, E represents the end point of the signal transient, b represents the median of the transients duration of some number of records, R is an integer random number

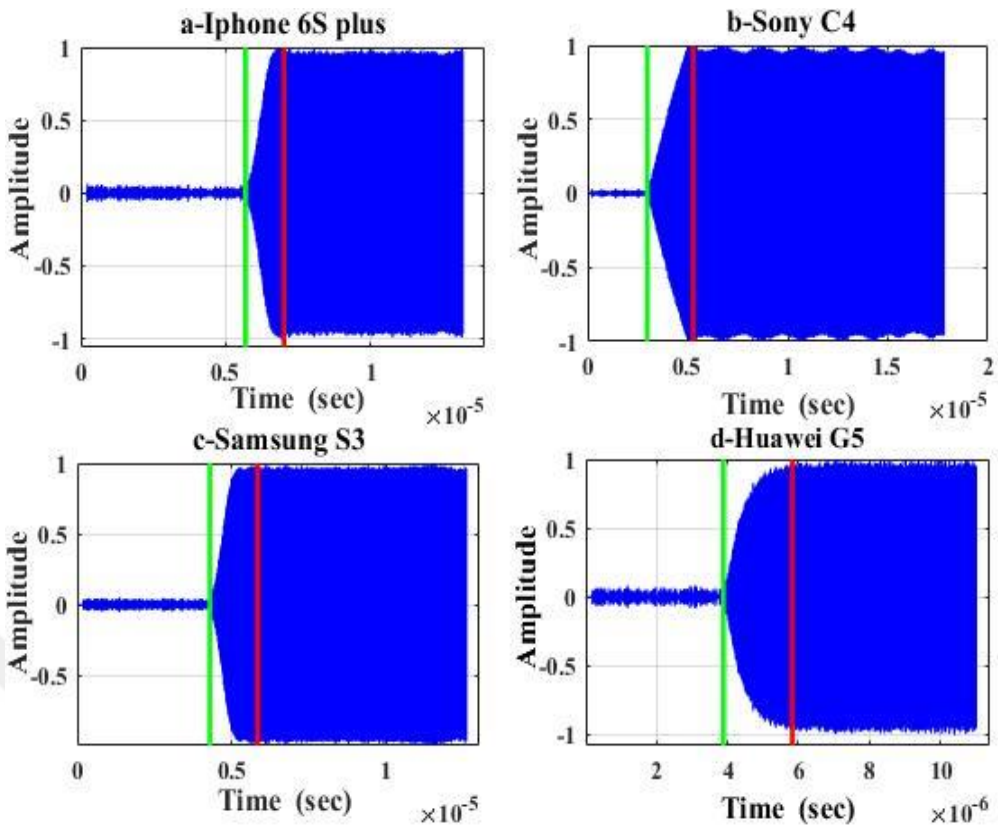


Figure 5.5 The detected transients of BT signals

5.3 Obtaining the Hilbert spectrum

After detecting the transient signal, HHT will be applied to them in order to generate its Hilbert spectrum.

The discrete Hilbert transform (DHT) is used to compute the analytic signal for a real valued signal [27, 28], where the analytic Signal is define as a complex signal whose imaginary part is the Hilbert transform of its real part as discuced in section 3.2. The transmission data represents a sequence of nonlinear and non-stationary real valued data. In order to analyze this complicated data set, we need to decompose it to sub-data sets each of which called intrinsic mode function (IMF) as discussed in section 3.1, which have well-behaved Hilbert transforms. The decomposition of the transmission data can be accomplished by EMD.

Obtaining the transients of the considered 20 cell phones, the HHT is applied to the detected transients' signals based on the HS procedure that given by chapter 3. The HHT is applied IMFs of the transient signals, where only significant IMFs are

considered. In this study, the HHT is applied to the three first high frequencies IMFs. That is because these IMFs have the highest amplitude values comparing to the rest of transient signal IMFs as shown in Fig. 5.6 and Fig. 5.7.

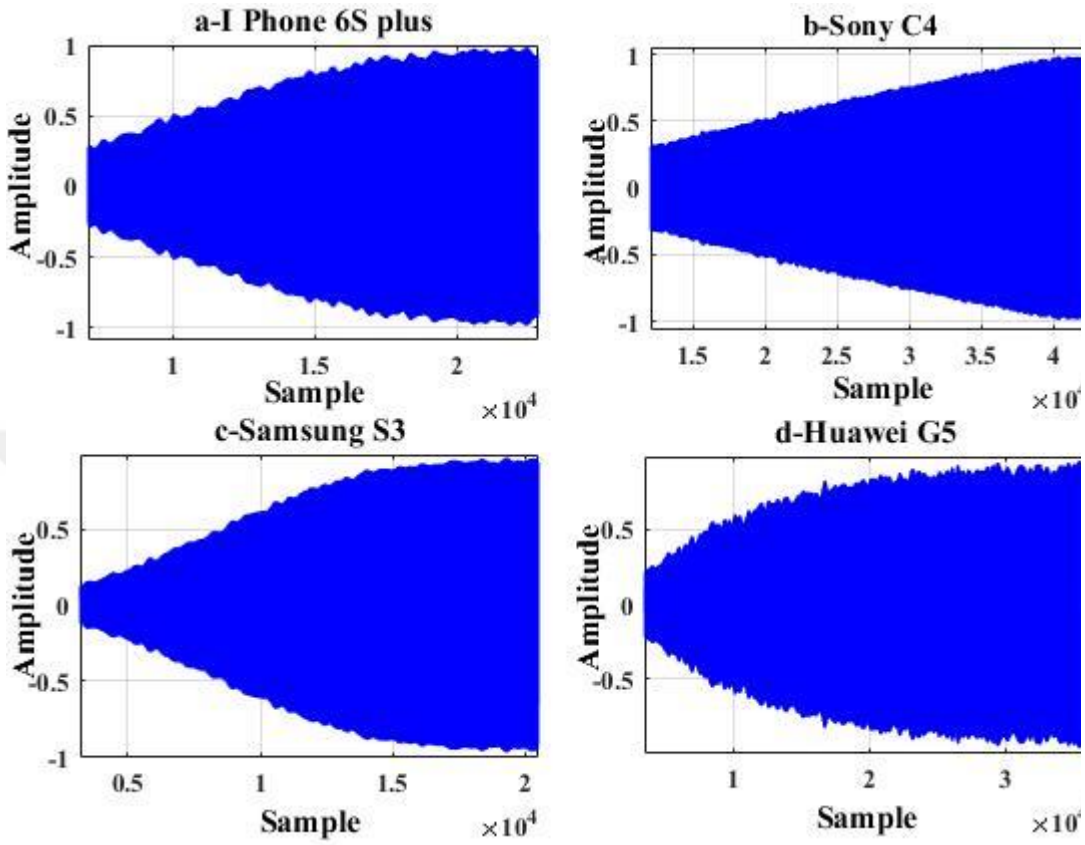


Figure 5.6 The first IMFs of the normalized transient signals.

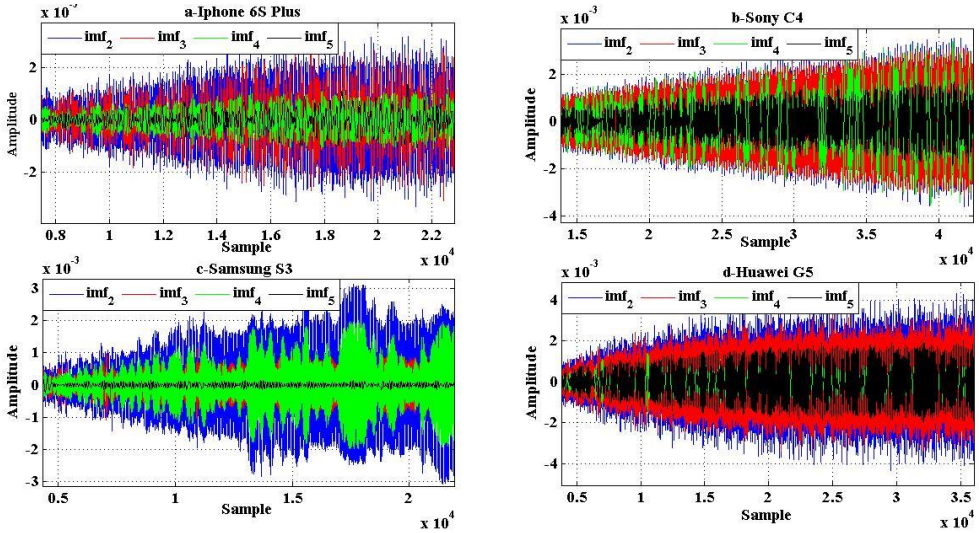


Figure 5.7. The 2nd through 5th IMFs of transient signals

The EMD method decomposes the IMFs based on their frequencies, the highest frequency the first decomposed. It's observed that the highest frequency imf has the highest amplitude as well. We can see from figure 2 that the amplitudes' values are in descent order from the 2nd imf through 5th imf. That why, in our work, we consider only the first three IMFs and neglect the rest. Consequently, this will reduce the computation cost and it will not affect the cumulative energy of the generated transient signal HS.

The HHT is implemented to the transients in order to generate subtle features of the BT phones' signals. As examples the Hilbert Spectrums of the transient signals that are illustrated in fig. (5.5) are shown in Fig. 5.8.

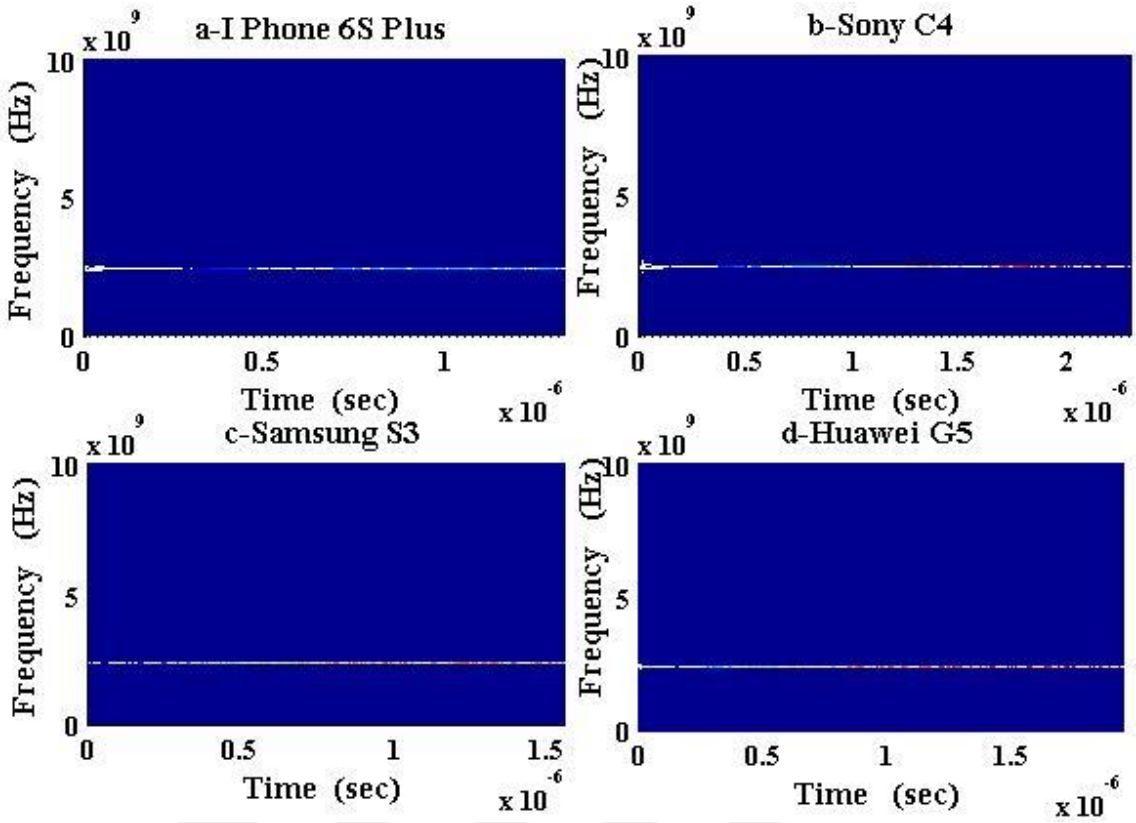


Figure 5.8 The TFEDs plane of transient signals

The BT transient signals HS of I-Phone 6s plus, and Sony-C4 are illustrated as time frequency energy distribution planes. It is clear from the fig. 5.8 that the two planes have the same frequency axes lengths whereas they have different time axes lengths. That is because duration of the BT transient signals are different based on the Phone brand, model and serial numbers. It's also inferred from the figure that the energy of the TFED are concentrated about 2.4 GHz. That is due to the range of the BT frequencies which are all located within the 2.4 GHz. It typically extends from 2.400 GHz to 2.4835 GHz. To obviously show the frequency range of the energy distributions in the TFEDs planes, the planes are zoomed around 2.4 GHz. The zoomed TEFDs are shown in Fig. 5.9.

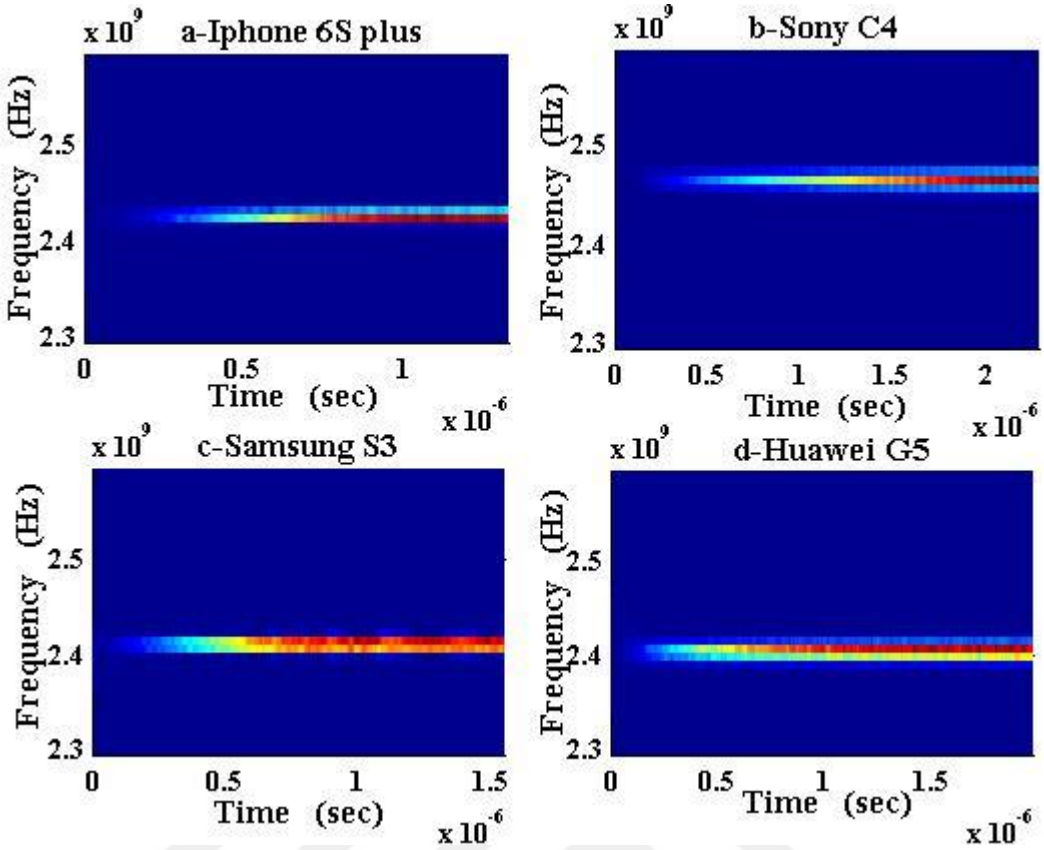


Figure 5.9. The zoomed TFEDs plane of transient signals

The frequencies of concentrated energy in TFEDs of the considered four BT signals are 2.427, 2.469, 2.41, and 2.41 for I-Phone 6s plus, Sony-C4, Samsung-S3, and Huawei-G5 signals, respectively. Based on our study the concentrated energy on the frequency axis in BT signals TFED cannot be a distinguishing features. That is because the BT signal frequency does not depend on the cell phone type. However the quantities of the energy in the signals are varying based on the cell phones brand, model, and serial number. These variations are demonstrated as three-dimension plots in fig. 5.10.

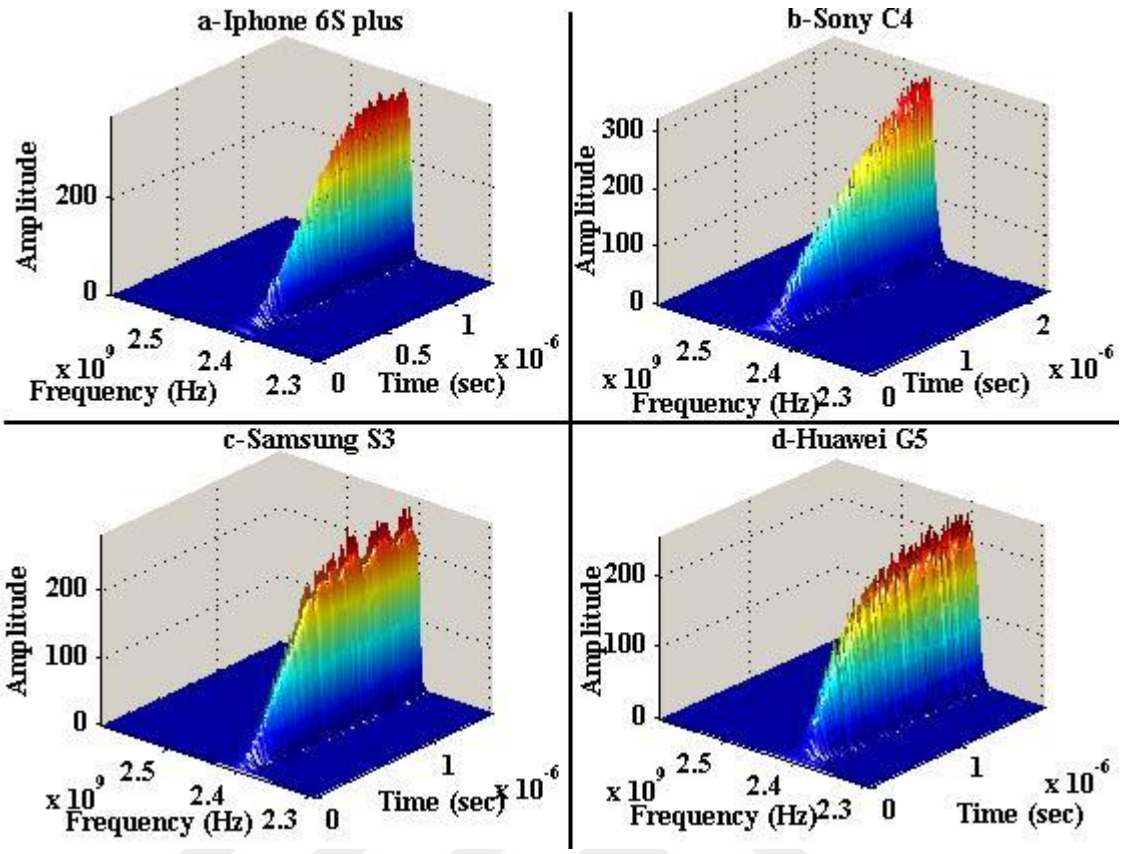


Figure 5.10. Three dimensional TFEDs of transient signals

The summation of energy of HSs that are shown in Fig. 5.10 can be computed along time and frequency axes. These summations generate two vectors of energy as shown in fig. 5.11. In which the summation of energy along time axis, of the four considered BT signals, is illustrated in Fig. 5.11-a, whereas their energy summation along frequency axis is shown in Fig. 5.11-b.

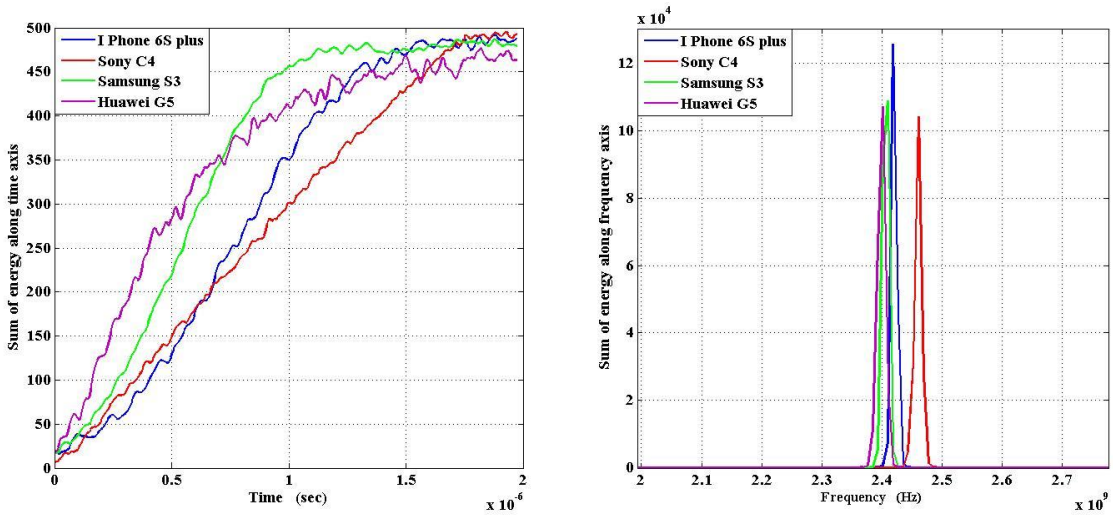


Figure 5.11. a- The summation of HS energy along time axis, b- The summation of HS energy along frequency axis.

Figure 5.11 represents samples of separated energy curves. Each of which represents just summation of energy of one record out of 150 records belongs to one cell phone out of 20 introduced cell phones. Twenty separated curves' bunches are generated. Each separated bunch composed of 150 energy curves.

It's inferred from fig.5.6 that the curves of the summation energies are varying, so many distinguishing features can be extracted from them. These features along with the overall features that are directly extracted from the transient signals are utilized to identify and classify the introduced cell phone devices. The maximum energy of the transient signals, the duration of transient signals, and the summation of the transient signals energy of the considered four cell phones are illustrated in table 5.2.

Table 5.2 A sample of some different features of four different cell phone transient signals.

Cell Phone Device	Max. E. of transient (mv)	Sum. E. of transient (mv)	Duration of transient
I-Phone 6S Plus	**0.0018**	**-0.0022**	**26526**
Sony C4	**0.0034**	**0.0023**	**45718**
Samsung S3	**0.0028**	**0.0037**	**31018**
Huawei G5	**0.0012**	**-0.0016**	**38778**

The illustrated features are extracted from transients of denoised BT signals. The summation of energy curves that is shown in Fig. 5.11, and the overall features which are demonstrated in table 5.2 give us an indications of the distinction and uniqueness of each cell phone device before extracting the rest of features. The overall features along with the subtle features are used to classify and identify the studied 20 different cell phones. This can be accomplished by extracting and preparing the features before applying the classifiers.

5.4 Extracting and smoothing of features

The BT transmitted signal contains a RF fingerprint. The signal RF fingerprint is associated with the cell phone brand, model and serial number. The RF fingerprint composed of bunch of unique features. The features that are used in this study are extracted from the transient signal and the HS of the transient signals as mentioned in the previous section. Thirteen features are extracted from each transient signals. These features are divided into three groups, namely, features directly extracted from the transient signals, features extracted from the summation of energy along time axis of the TFEDs of transients, and features extracted from the summation of energy along frequency axis of the TFEDs of transients as illustrated in Table 3.

Table 5.3. RF fingerprints

Features group	RF fingerprint	Feature label
Transient signal and energy envelope	**Duration of transient**	**f1**
	Summation of transient energy	**f2**
	Summation of transient energy envelope	**f3**
	Variance of transient energy envelope	**f4**
	Standard deviation of transient signal instantaneous phase	**f5**
	Entropy of transient signal instantaneous phase	**f6**
TFED of the transient signal a long time axis	**Length of summation of energy along time axis**	**f7**
	Slope of summation of energy along time axis	**f8**
	Variance of summation of energy along time axis	**f9**
	Maximum of summation of energy along time axis	**f10**
	Polyfit coefficient of summation of energy along time axis	**f11**
TFED of the transient Signal along frequency axis	**Maximum of summation of energy along frequency axis**	**f12**
	Variance of summation of energy along frequency axis	**f13**

More features can be extracted from the transient signals, transient instantaneous characteristics, and the transient TFEDs. The extracted features that are listed in Table 5.3 represented the most robust features by which the RF fingerprint can be generated.

In order to show the robustness of extracted features before applying the classifiers, two types of plots are used, namely, box plot and conventional two dimensional Matlab plot. The box plot technique is used to illustrate the extracted features. This plot (see Fig.5.12) is defined by means of the following five criteria: The lower adjacent, the 25th percentile, the median, the 75th percentile, and the upper adjacent. The 25th percentile is a value of the considered data group at which 25% of the data values are less than it. A 50% of the data group or the Inter-Quartile Range (IQR) are located between the 25th percentile and the 75th percentile. This means that 25th of the data values are greater than the 75th percentile value. The median criterion which is illustrated by a red line in the box middle represents the 50 percentile.

The lower and upper adjacent represent the whisker boundaries. These boundaries can be determined by multiplying a value (usually 1.5) times the IQR, and subtracting or adding it to the 25th percentile or the 75th percentile, respectively. The

values of the data group that are out of the whisker boundaries are outlier values [45].

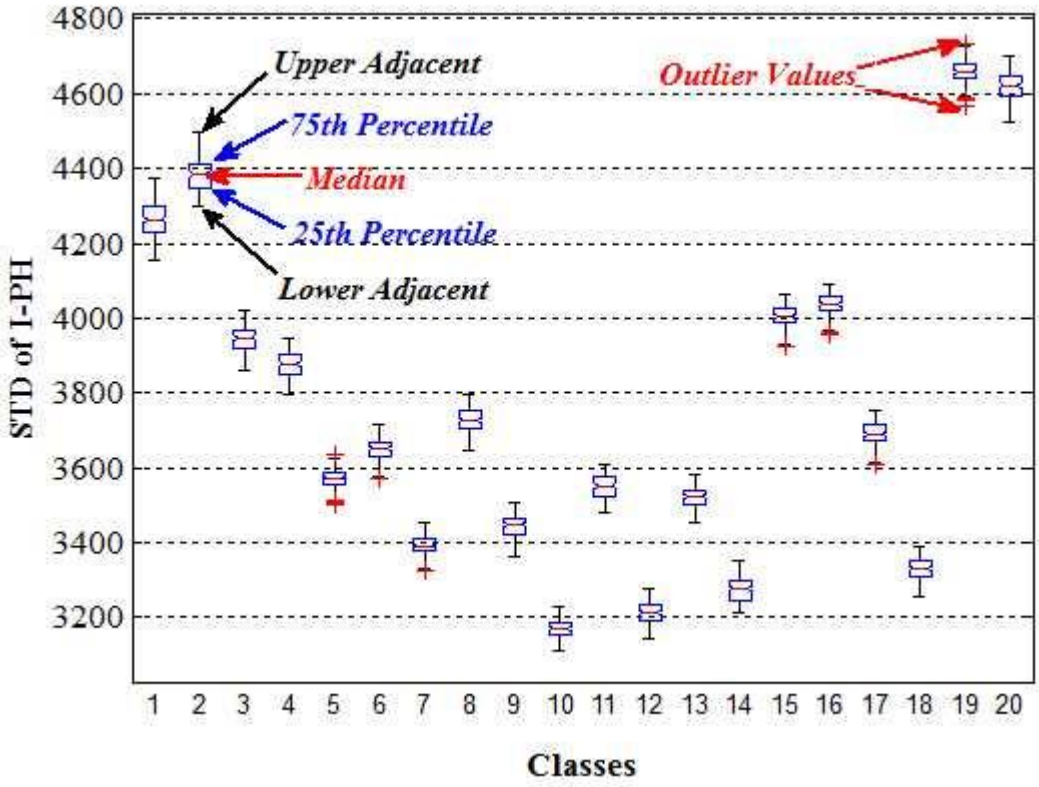


Figure 5.12 The box plot criteria

Because of the separation property, the standard deviation (STD) of transient signal instantaneous phase (f5) is one of the most robust features. This feature is extracted from the instantaneous characteristics of the transient signal. The STD of transient signals' instantaneous phases of the twenty introduced classes with 150 records is introduced as an example of the robust features. The smoothing of this feature is shown in Fig. 5.13 through Fig. 5.18. Fig. 5.13 shows a conventional two dimensional Matlab plot of the (STD) of transient signals' instantaneous phases. It's inferred from Fig. 5.13 that some classes are completely separated. For instant, class_1 (Huawie G5 A) is completely separated from class_3 through class_20, however class_1 is partially separated (nearly separable) from class_2 (Huawei G5 A). Some other classes are nearly separable such as class_9 (Iphone_7_A) and class_11 (LG_G4_B). However, there are some classes are barely separated such as class_15 (SamSung_S3_A) and class_16 (SamSung_S3_B). This deficiency can be

remedied by feature smoothing can be compensated by other features in the RF fingerprints hyperspace. The advantage of the conventional two dimensional Matlab plot is that by which we can know the number of interfered records and interfered records' numbers for each class.

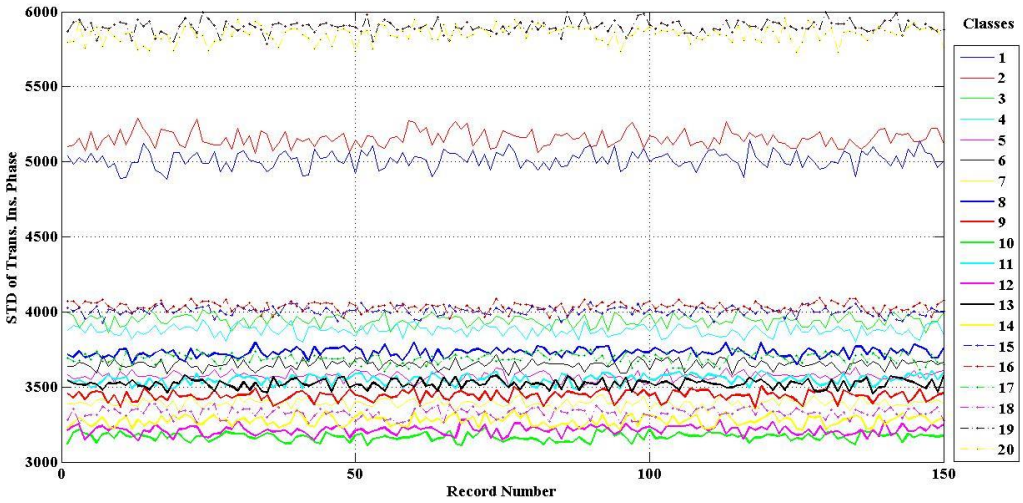


Figure 5.13. A conventional plot of STD of trans. ins. phases of 20 cell phones

The widely used box plot is implemented to the STD of transient signals' instantaneous phases of the twenty introduced classes. That is illustrated in Fig. 5.14. By means of the box plot a clear picture of the interference among classes feature can be seen. Moreover, by using the box plot, the percentage of the data concentration can be recognized.

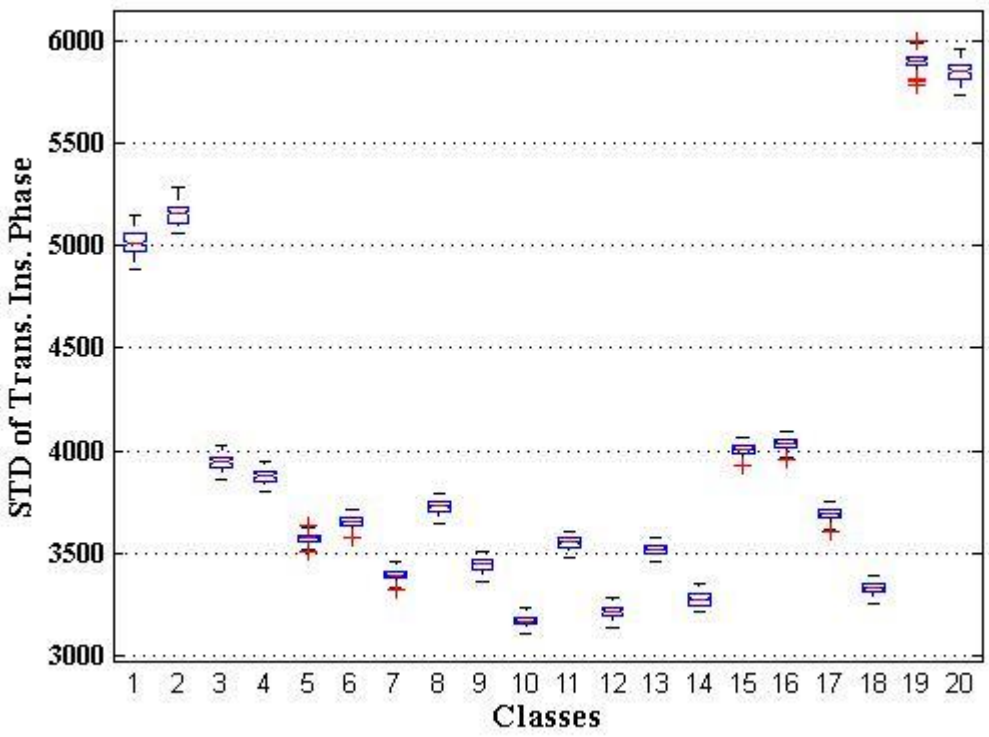


Figure 5.14. The box plot of STD of trans. ins. phases of 20 cell phones

In order to clearly recognize the interference among classes feature and enhance the robustness of the features, a smoothing process is needed. The smoothing process that is considered in this study starts with the data zooming. The data zooming is accomplished by making the overall range of the feature data small as possible. For instant, it's obvious from Fig. 5.13 and Fig. 5.14 that the majority of data is located between 3000 and 4000. However, there are four classes' features' data, namely, class_1, class_2, class_19 (Sony_C4_A), and class_20 (Sony_C4_B) are located extremely far from the mean of the majority of the data. By multiplying the extreme data by certain factors the all data will concentrate in the possible smallest range. The zooming of the considered STD of transient signals' instantaneous phases feature is shown in Fig. 5.15 and Fig. 5.16.

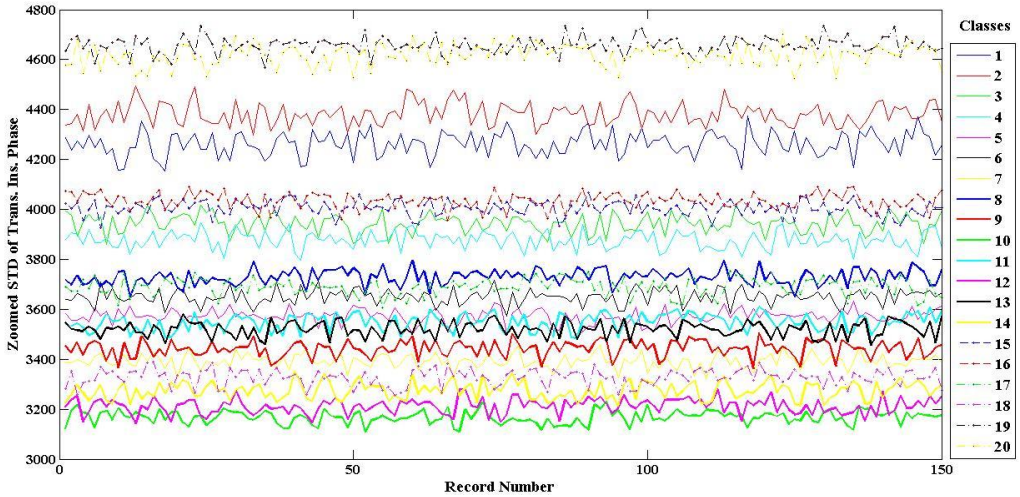


Figure 5.15 Zooming of the STD of transients' instantaneous phases

The comparison between Fig. 5.13 and Fig. 5.15 demonstrates that the discrimination of the feature curves becomes more obvious after data zooming implementation. Moreover it becomes easier to recognize the number of the interfered records and the interfered records numbers. Fig. 5.16 shows the zooming of Fig. 5.14. The boxes of the data become larger and clear after the zooming process. In addition, the upper adjacent and outer values can be seen clearly.

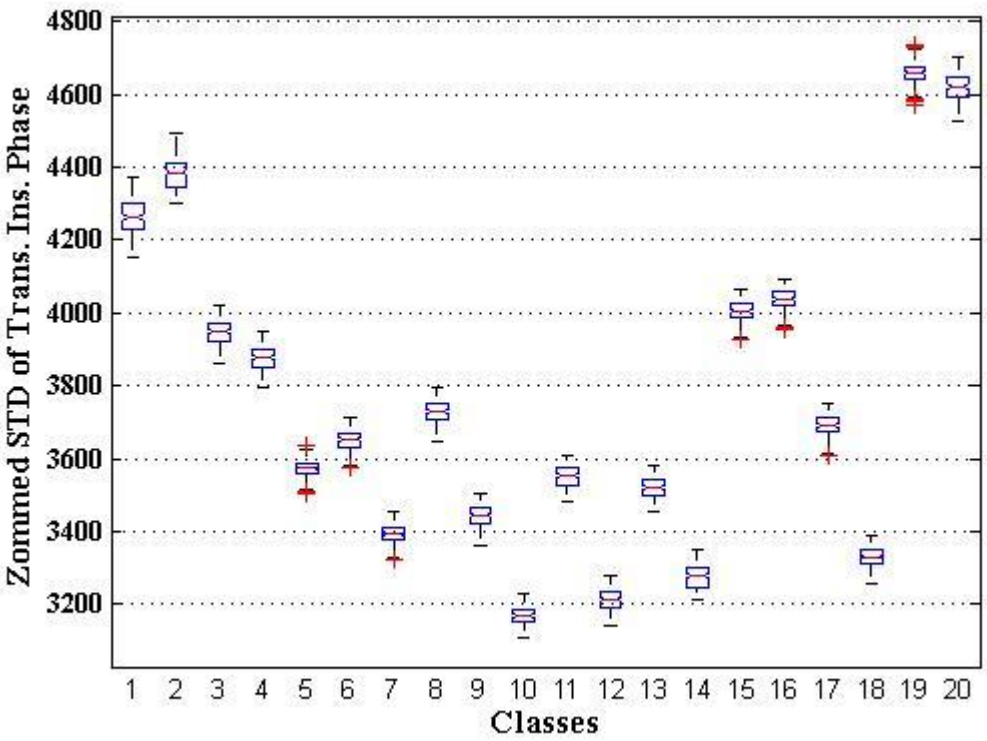


Figure 5.16 Box plot of zoomed STD of transients' instantaneous phases

The zooming process generates clear features' plots; however, it doesn't enhance the robustness of the features. In order to enhance the features robustness, the variance of each class feature vector should be decreased. This can be done by applying a filtering or smoothing technique to the extracted zoomed features. The parameters or the coefficients of the applied filter must be the same for all data of classes considered features, otherwise this will corrupt the whole identification process. The median filter is used to smooth some of extracted features. As an example of smoothed features, the STD of transients' instantaneous phases feature is smoothed by means of median filter. The interference among classes feature is decreased after filtering process as shown in Fig. 5.17 and Fig. 5.18.

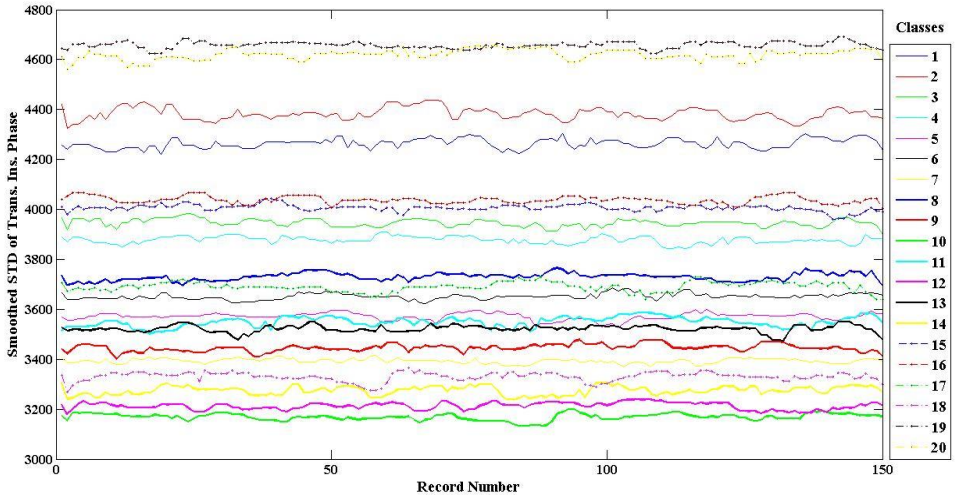


Figure 5.17 Smoothing of the STD of transients' instantaneous phases

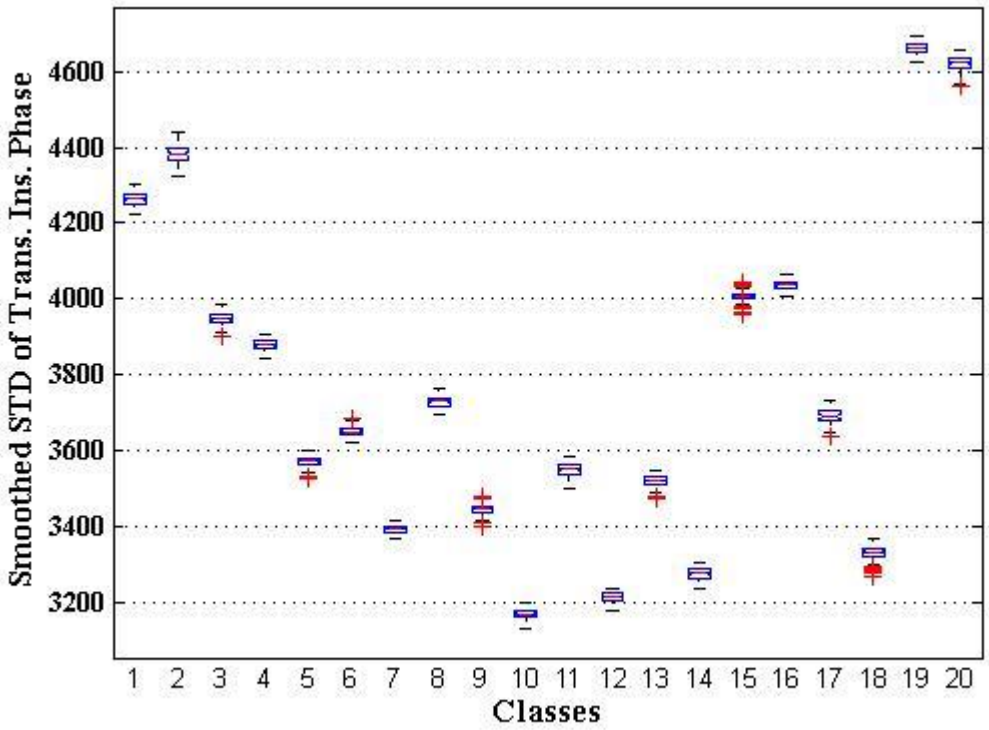


Figure 5.18 Box plot of smoothed STD of transients' instantaneous phases

It's concluded from Fig. 5.17 that the data range of each vector of class features becomes narrow. It's also obviously recognized that the features boxes in Fig. 5.18 become smaller than the boxes in Fig. 5.16. That means the classes features

interferences are decreased. Six classes are chosen to clearly illustrate the effect of smoothing process. The classes are class_1, class_2, class_15, class_16, class_19, and class_20 which are Huawei G5 A, Huawei G5 B, Samsung S3 A, Samsung S3 A, Sony C4 A, Sony C4 B, respectively. It is notice that the interference occurs between the same brands same models different serial numbers classes. This interference is decreased after applying the median filter to the unsmoothed features' vectors as shown in Fig. 5.19.

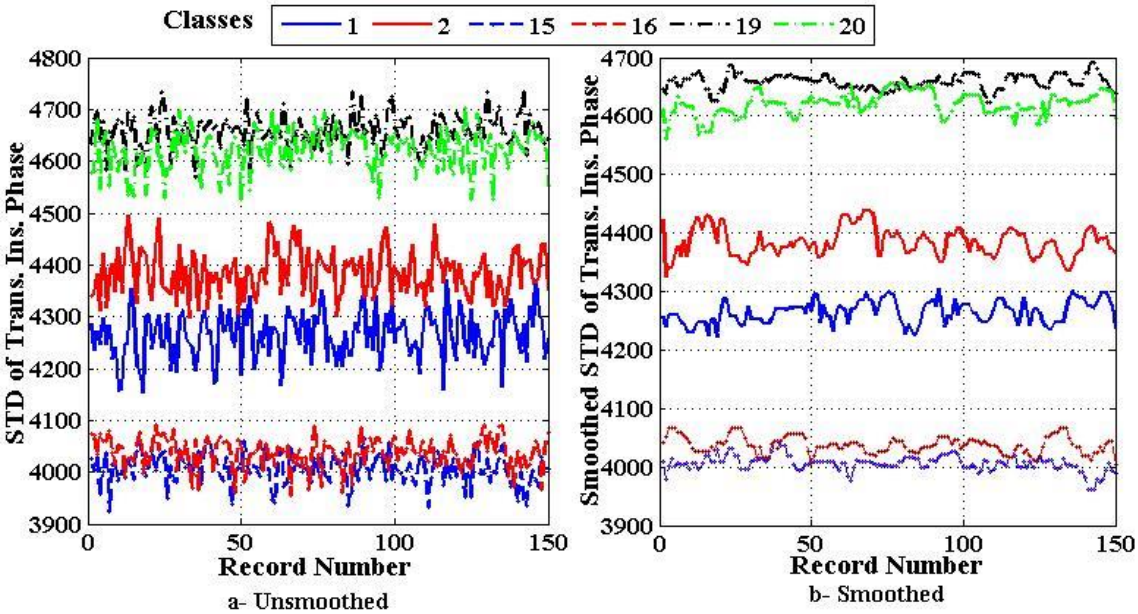


Figure 5.19 Unsmoothed vs. smoothed STD of six classes trans. ins. phase

In order to enhance the robustness of the RF fingerprint, the smoothing process is applied to some other considered features. As aforementioned, even though the smoothing process decreases the interference among classes' features nevertheless there are still partially and barely separable classes. This feature deficiency can be compensated inside the RF fingerprint hyper space, such that, the partially separated or barely separated classes based on a certain feature are completely separated when another feature is considered.

Based on STD of transient instantaneous phase, class_3 (I phone 5S A), and Class_4 (I phone 5S B) are partially separable (see Fig. 5.12, and Fig. 5.13). This feature deficiency is compensated by the entropy of transient instantaneous phase. It is

obvious from Fig. 5.20 that the class_3 and class_4 are completely separable, and vice versa. For instant, based on entropy of transient instantaneous phase, class_11 (LG G4 A), and Class_11 (LG G4 B) are barely separable (see Fig. 5.20). This feature deficiency is compensated by the STD of transient instantaneous phase. It is obvious from Fig. 5.18 that the class_11 and class_12 are completely separable

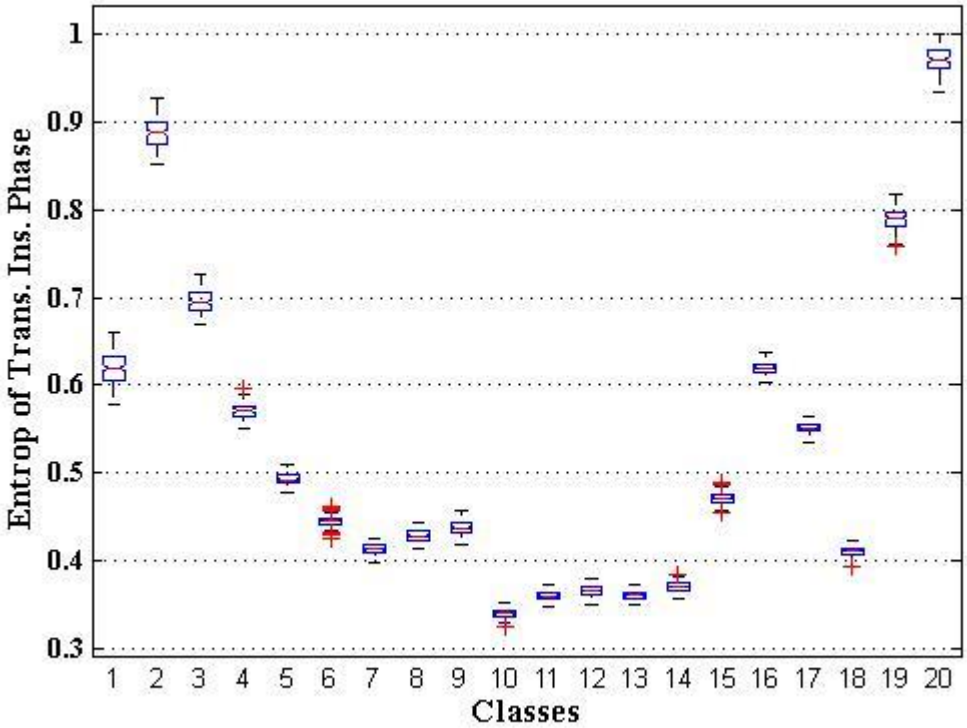


Figure 5.20 Box plot of smoothed entropy of transients' instantaneous phase

In general the classification and identification of classes are accomplished by the contribution of all features in the RF fingerprint hyperspace. The feasibility and the robustness of the extracted features can be measured before applying the classifiers. In order to examine the features robustness, the features distribution of two or more classes must be investigated. Same-brands same-models and different serial numbers cell phones can be distinguished by using some robust features in the RF fingerprint (see Fig. 5.21). The considered two cell phones, in this case study, are Samsung Note3 A and Samsung Note3 B, which are defined by class_13 and the class_14, respectively. Due to the completely separable features f1, & f5, and the nearly separable f7, f9, & f11 these two classes can be perfectly distinguished.

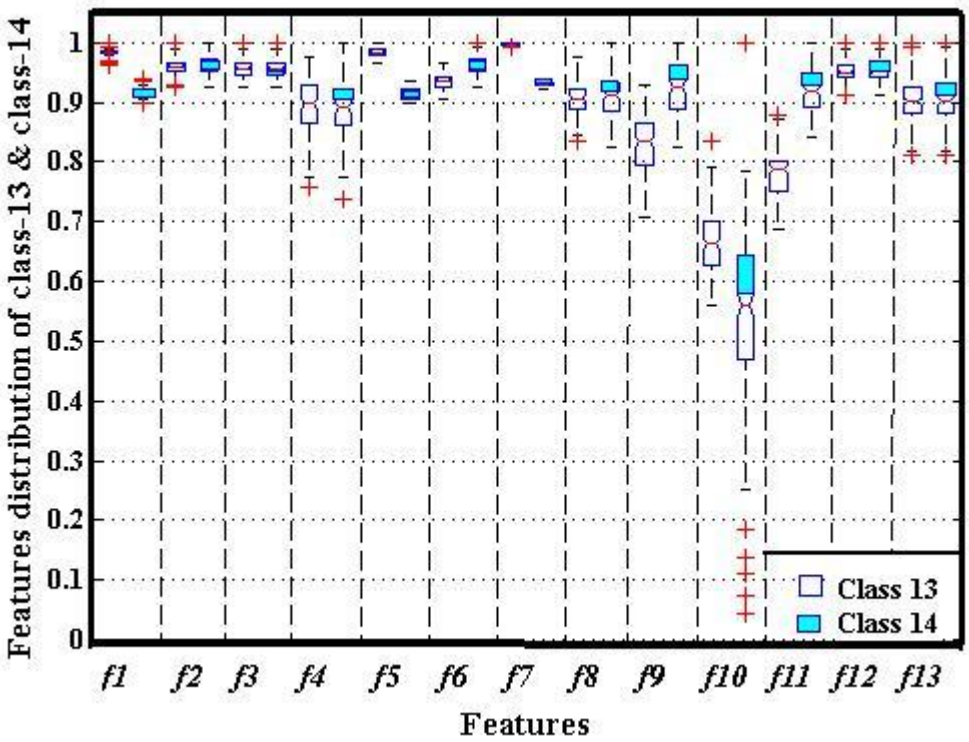


Figure 5.21 Features distribution of class-13 & class-14

Two same brand different model cell phones, namely class_8 (I Phone 6S Plus B), and class_9 (I Phone 7 A) are subjected to different features (see Fig. 5.15). These two classes have the most interfered features among the considered classes in our study. It's reasonable to deduce that the introduced features can distinguish the class_8 and the class_9 perfectly. That is owing to the nearly separable property of features f7; f8; & f10; and the completely separable property of features f1; & f5.

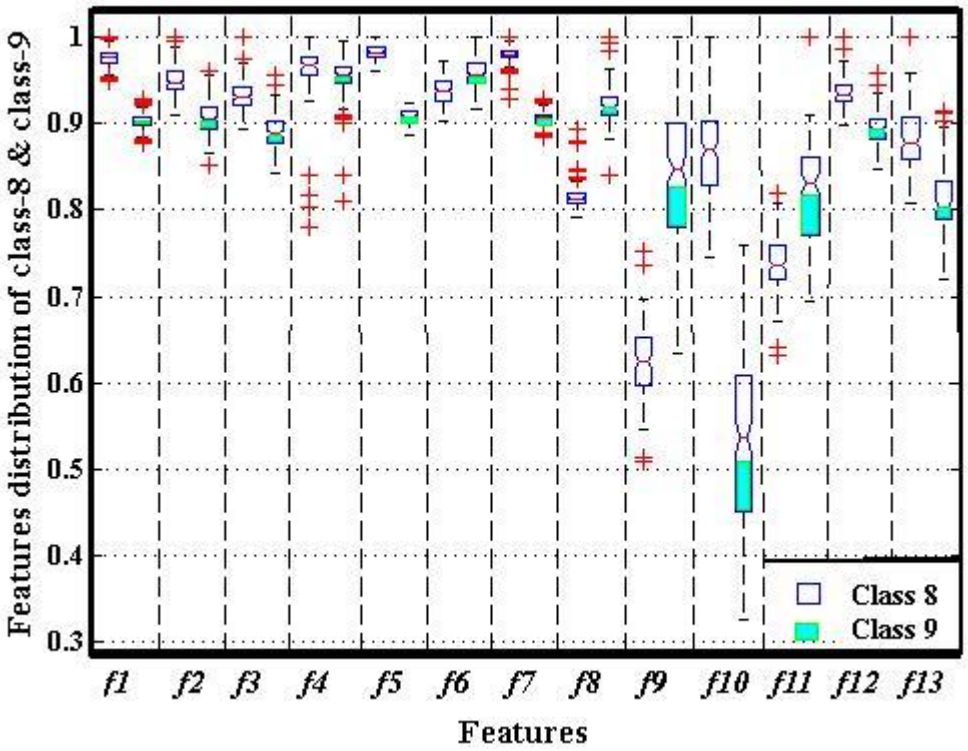


Figure 5.22 Features distribution of class-8 & class-9

Each of the rest of the twenty studied classes is classified and identified based on its extracted unique features. Based on the BT signals' records extracted features of the studied cell phones, the classification of these devices can be accomplished.

5.5 Classification based on high SNR level

The introduced cell phone devices are classified and identified. The RF fingerprint which composed of thirteen features is used as input data to the classification process. Numbers of classifiers are applied, in this study, to accomplish the classification and identification of twenty deferent classes. Some classifiers are applied to the extracted features. A brief definitions and information about the Complex decision tree, Linear Support Vector Machine (SVM), and Linear Discriminant Analysis (LDA) classifiers are introduced in the following paragraphs.

The decision tree is a supervised learning algorithm that mostly used in classification problems. The tree is composed of three types of nodes; root node which contains all

the data, splits nodes which generate internal nodes of the tree, leaves which defined as terminal nodes. The splits nodes are estimated from a statistical procedure applied to the training data. The splits nodes are used as a mapping from the root node to the leaf decision nodes. The main idea of this procedure is splitting the data into subsets; each subset belongs to only one class. The procedure of the decision tree classifiers depends on recursively partitions of the input data set into smaller subdivisions. The learning algorithms of the decision tree require high-quality training data to generate the model. Therefore, a set of training data representative of the main data must be available to generate an accurate decision tree classifier model [46].

Support vector machine is a supervised machine learning algorithm that is utilized for classification and regression process. The idea behind the SVM procedure is mapping the input data x into high dimensional space of data by using a mapping function

$\phi(.)$ and linear function $f(x) = w_(x) + b$, where w, and b are optimized coefficients. The linear function is used to fit the data in the space and generate the hyperplane [47]. The construction of SVM is to minimize the upper bound error by maximize the margin between the separating hyperplanes and the separated data. The advantage of SVM technique is the property of condensing information in the training data by using small number of data points [47].

The RF fingerprint which is composed of the extracted features represents subjected data of the classifiers. These data is divided into two groups, namely, train data and test data. The classifiers are firstly trained and then applied to the test data. The training stage of each classification process generates model. The generated model is used to classify the classes based on the test data. Table 4 shows the accuracies and performances of different classifiers applied to unsmoothed RF fingerprint. The accuracies are obtained in both training and testing stages of the classification processes. The data is divided as 40 % train data and 60% test data.

Table 5.4 classifiers' accuracies based on 40%-train 60%-test unsmoothed RF fingerprint.

Classifier	Accuracy of training stage (%)	Performance of testing stage (%)
Complex Tree	92.0	97.9
Medium Tree	84.1	92.2
Linear Discriminant	97.3	91.3
Quadratic Discriminant	95.0	94.2
Linear SVM	98.9	97.2
Quadratic SVM	97.1	97.2
Fine KNN	89.0	90.1
Cubic KNN	86.1	84.9
Ensemble Bagged Trees	98.3	92.4
Ensemble Subspace Disc.	95.7	93.7

In order to investigate the effect of the data division size (the percent of train data to the percent of test data), the same unsmoothed RF fingerprint are divided into 20% train data and 80% test data. The introduced classifiers are applied to the 20%-train 80%-test unsmoothed RF fingerprint. The classification train accuracies and test performances are illustrated in the following table:

Table 5.5 classifiers' accuracies based on 20%-train 80%-test unsmoothed RF fingerprint.

Classifier	Accuracy of training stage (%)	Performance of testing stage (%)
Complex Tree	95.2	97
Medium Tree	93.7	95.7
Linear Discriminant	91.0	91.1
Quadratic Discriminant	95.3	92.7
Linear SVM	95.5	95.4
Quadratic SVM	94.8	94.6
Fine KNN	87.0	85.3
Cubic KNN	85.2	80.1
Ensemble Bagged Trees	96.2	97.6
Ensemble Subspace Disc.	95.3	93.7

The average test performance of classifiers based on 40%-train 60%-test unsmoothed RF fingerprint is 93.11%, whereas, the average test performance of classifiers based on 20%-train 80%-test unsmoothed RF fingerprint is 92.32%. It's concluded from Table 5.4 and Table 5.5 that the test performances are slightly increased as the train data percent increases. The robustness and the feasibility of the generated RF fingerprint are enhanced in the features smoothing stage. The median filter is used to smooth the vector of each feature fingerprint. In order to investigate the effectiveness of the features smoothing stage, the considered classifiers are applied to the smoothed RF fingerprint. The classifiers' accuracies and performances are evaluated based on 40%-train 60%-test smoothed RF fingerprint. The comparison among scatter plots of the train data are shown in Fig. 5.23 through Fig. 5.26. The unsmoothed features duration of transient (*f1*) against summation of transient energy (*f2*) are plotted as scatter plot for 20 classes (see Fig. 5.23). Applying smoothing

process the scatter plot of the same features is shown Fig. 5.25. The scatter plots of unsmoothed and smoothed features *f1* and slope of summation of energy along time axis of TFED (*f8*) are shown in Fig. 5.24 and Fig. 5.26, respectively.

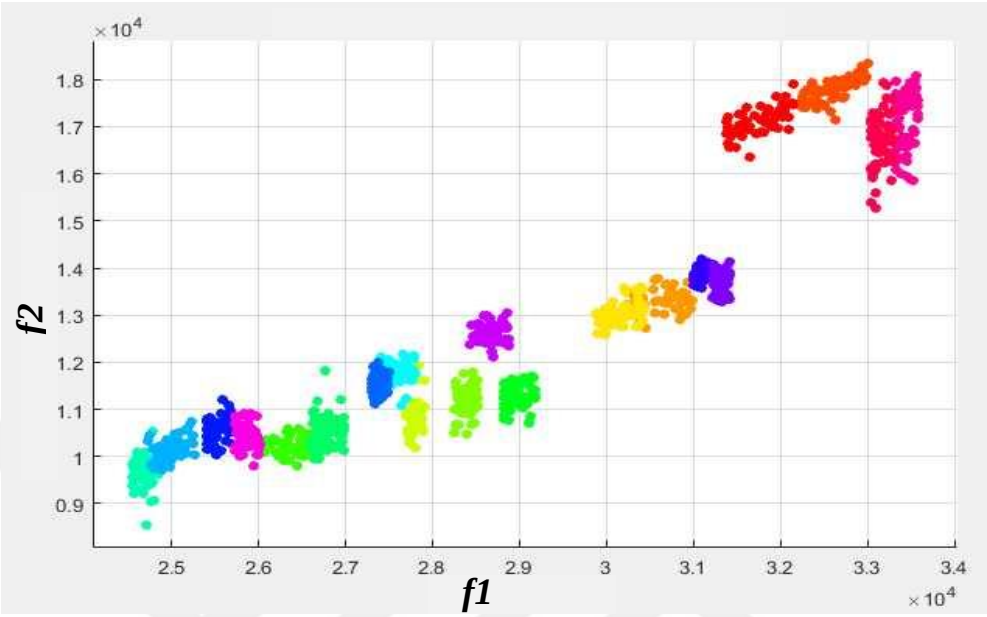


Figure 5.23 Scatter plot of unsmoothed features f1 vs . f2

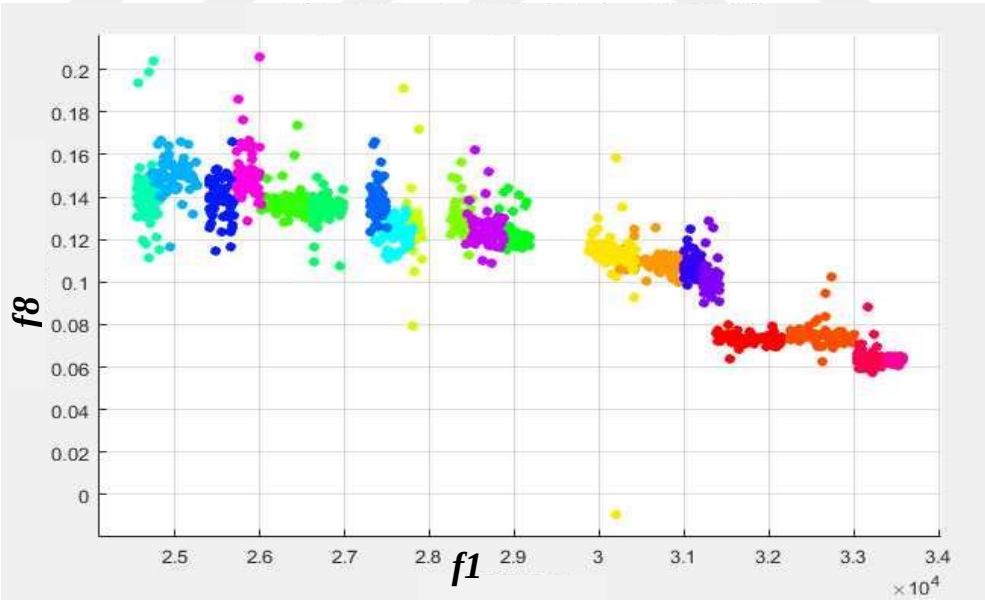


Figure 5.24 Scatter plot of unsmoothed features f1 vs. f8

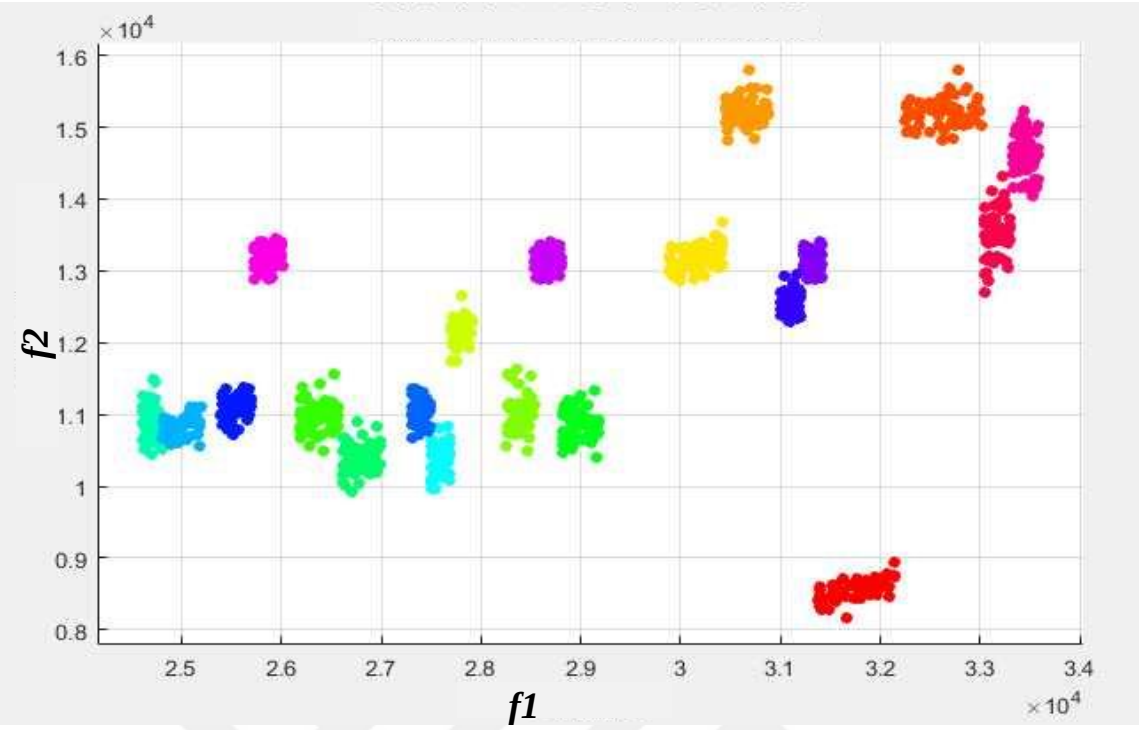


Figure 5.25 Scatter plot of smoothed features f1 vs. f2

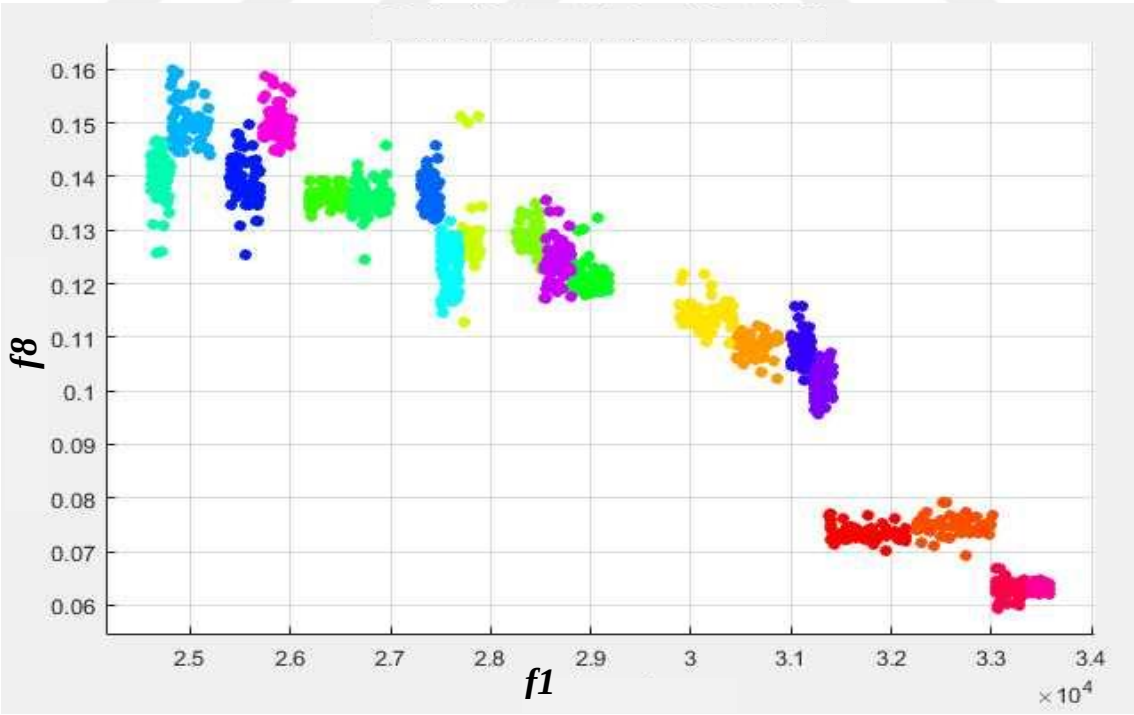


Figure 5.26 Scatter plot of smoothed features f1 vs. f8

It's deduced from the features scatter plots that the features' plots of each class records are close to each other after smoothing process. That is because the median filter which is applied in the data smoothing gathers each window data around their median.

A certain classifier is applied (as an exampl) to train data to generate a classifier model. The generated model is utilized to classify the test data based on the classifier training or learning. The complex tree training confusion matrices based on 40%-train of unsmoothed and smoothed RF fingerprints are shown in Fig. 5.27 and Fig. 5.28, respectively. The training confusion matrices demonstrate 98.3% and 100% for unsmoothed and smoothed features respectivily.

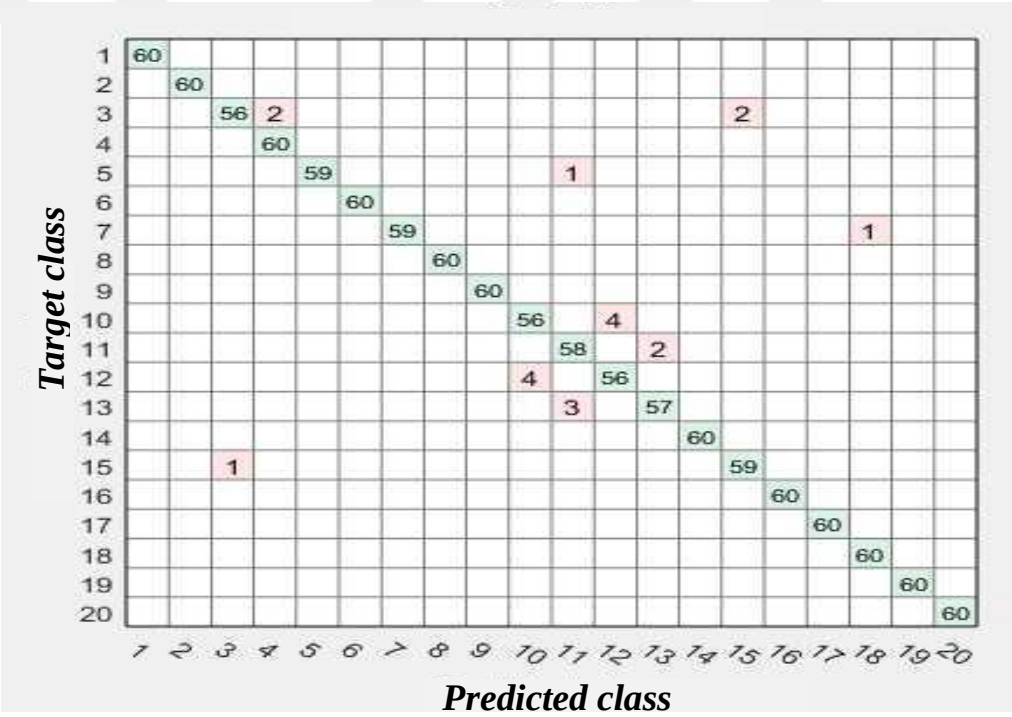


Figure 5.27 Complex tree classifier confusion matrix based on 40%-train of unsmoothed RF fingerprint

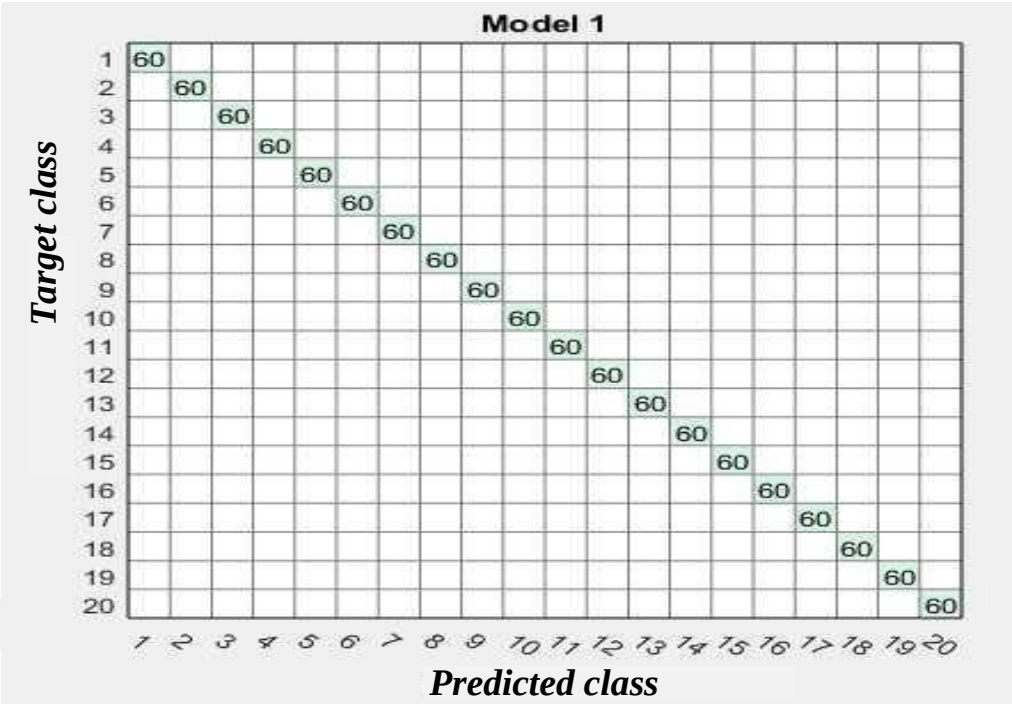


Figure 5.28 Complex tree classifier confusion matrix based on 40%-train of smoothed RF fingerprint

It's concluded from Fig.5.23 through Fig. 5.28 that the scatter plots along with confusion matrices give a prediction about the classification performance before applying the test classification. Each classifier among the introduced classifiers generates a training confusion matrix and classifier model. The classifier model is used to classify the classes based on the test data. The considered classifiers that are applied to the unsmoothed RF finger print are applied after the data smoothing process. Table 5.6 demonstrates the classifiers' accuracies based on 40%-train 60%-test smoothed RF fingerprint.

Table 5.6 Classifiers' accuracies based on 40%-train 60%-test smoothed RF fingerprint.

Classifier	Accuracy of training stage (%)	Performance of testing stage (%)
Complex Tree	100	99.6
Medium Tree	100	99.6
Linear Discriminant	99.8	99.7
Quadratic Discriminant	99.9	99.9
Linear SVM	100	99.8
Quadratic SVM	100	99.9
Fine KNN	99.8	99.4
Cubic KNN	99.1	99.1
Ensemble Bagged Trees	99.9	100
Ensemble Subspace Disc.	100	99.8

It's inferred from table 5.6 that the classification process demonstrates high classification train accuracies and test performance . That is due to the convergence of all records feature of each class and the divergence of the deferent classes' features after the smoothing process.

In comparison between the classification performanves of classifiers based on unsmoothed and smoothed features, it's concluded that the classifiers' performances are increased after applying the smoothing process. The average performances of the ten applied classifiers based on smoothed features are 99.85% as train accuracy and 99.72% as test performance, whereas, they demonstrate, based on unsmoothed features, 93.19% as train accuracy and 93.07% as test performance.

As a result, the classifiers' accuracies based on 40%-train 60%-test are enhanced after the smoothing features process, so that both train and test classifiers' accuracies are improved by 7.15%.

As one of robust classifiers at low and high SNR levels, L-SVM is trained with train data (40%) and its training model is applied to 60% test data. The test confusion matrix of L-SVM classifier is shown in Fig. 5.29. The average test performance of L-SVM for all classes is 97.2 %, where four classes are 100% true classified.

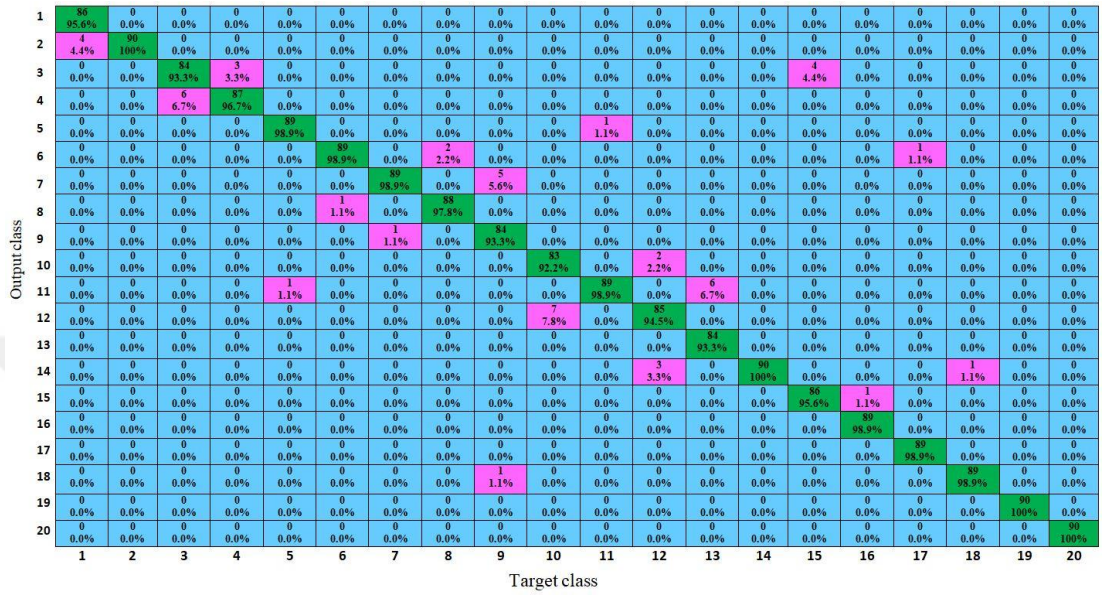


Figure 5.29 Test confusion matrix for 30-35 high SNR un-smoothed transients via L-SVM.

The performance of the classifiers are improved by smooth the features vectors before classification process. The test confusion matrix that is obtained by implementing L-SVM to smoothed 60 % features data is shown in Fig. 5.30. The test performance of L-SVM is improved by 2.675% after feature smoothing. Only three misclassified records were identified when smoothed features are used. In this case, the misclassified records are as follows: class 10 is misclassified as class 12 in one record, and class 8 is misclassified as class 6 in two records. This demonstrates that the use of the smoothed features increases the classification performance substantially.

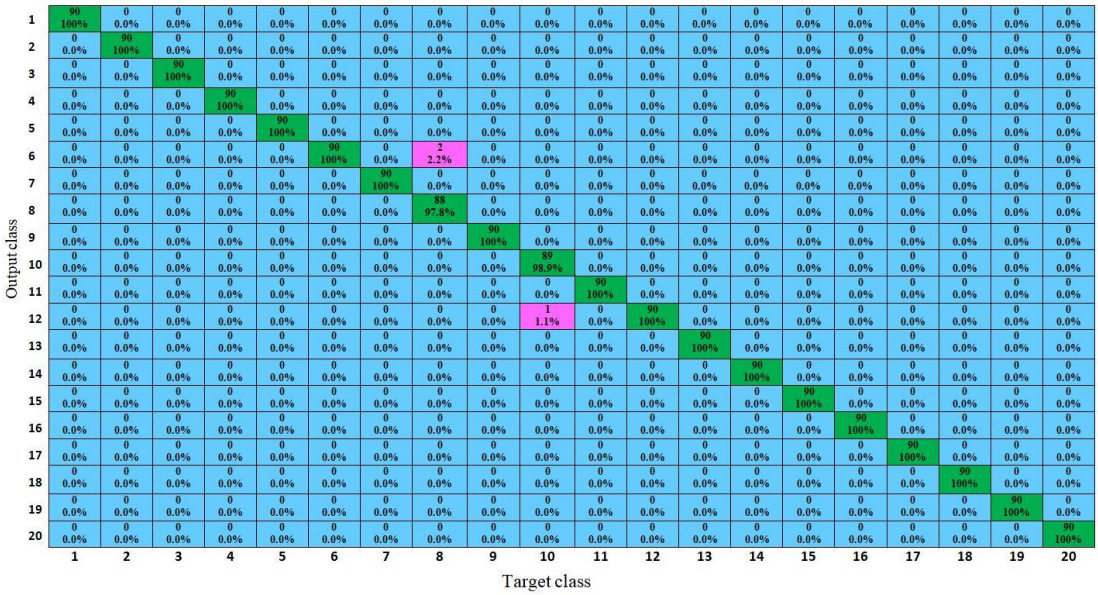


Figure 5.30 Test confusion matrix for 30-35 high SNR smoothed transients via L-SVM.

Two more classifiers along with L-SVM namly C-Tree and LDA are chosen to be used in the annlysis of noise effects on the classification performance. Table 5.7 illustrates the considered classifiers performances for smothed and unsmoothed data at high SNR levels (30-35).

Table 5.7 Classifiers' performances at high SNR (30-35dB)

Classifier	Smoothed	Unsmoothed
L-SVM	99.8%	97.2%
Complex Tree	99.6%	97.59%
LDA	99.7%	91.3%

The performance in Table 5.7 is obtained from training the classifiers with 40% train data and testing them by 60% test data. The classifiers performances that are illustrated in Table 5.7 are considered in the noise performance analysis. Three different levels of transient signals SNR lower than 30-35dB is studied in the next chapter to evaluate the effects of noise on the classifier performance. The noise performance is done based on 40% to 60% train and test data, respectively.

CHAPTER 6

NOISE PERFORMANCE ANALYSIS

In the previous chapter, the classification process performances are investigated via different classifiers for high SNR transient signals. In order to evaluate the effectiveness of the introduced successive signal processes, the classification performances are examined for different levels of SNR transient signals. The SNR levels of the transient signals are reduced by adding the recorded channel noise to the corresponding transient signals as shown in Fig_6.1. The noise portion is weighted and added to the transient signal for each record of each class. The weight factors are varied such that the SNR levels of noisy transients are classified into three levels. These levels are 8-10 dB, 12-15 dB, and 18-23 dB. All classes' transients are considered for each SNR level. The SNR of the noisy transient are calculated based on the following equation:

$$SNR = 10log_{10}\left(\frac{s}{n} - 1\right) \tag{6.1}$$

The symmetric distribution of the noisy transient SNR levels is generated to match the Gaussian distribution (see the histogram in Fig. 6.2), so that the median of noisy transients SNR approximately equal its mean. This guarantees an appropriate distribution of the data SNR of all SNR levels. Consequently, sensible results can be achieved.

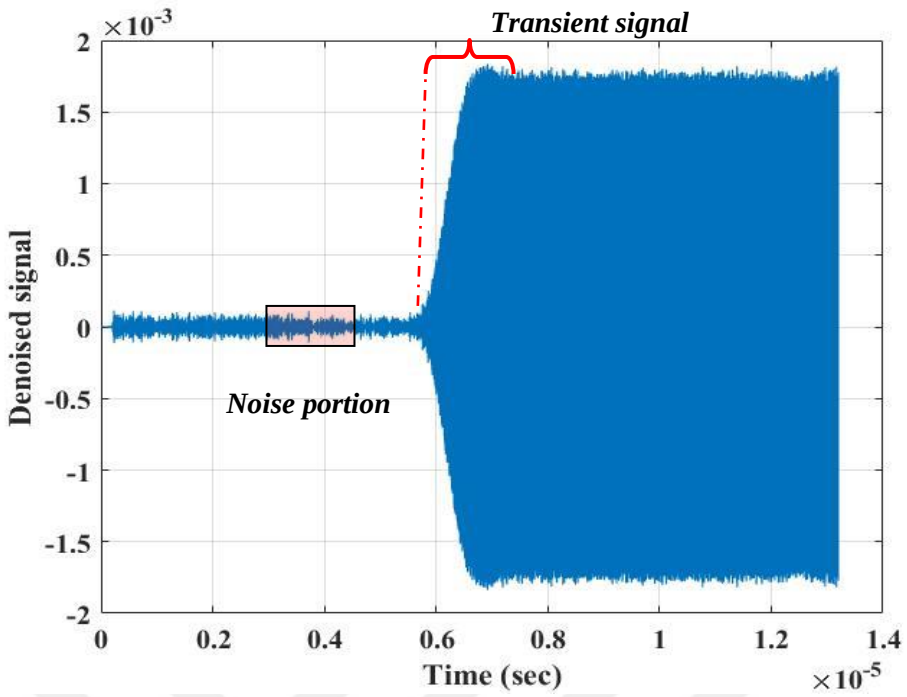


Figure 6.1 A recorded channel noise and associated transient signal.

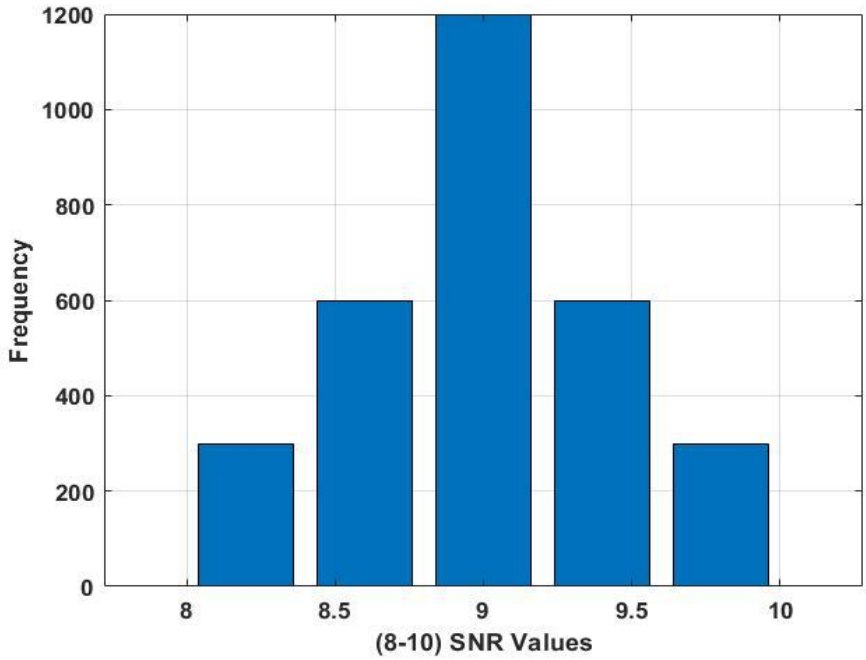


Figure 6.2 The Histogram of the SNR level values of 3000 noisy transient records.

Once the three levels of noisy transient are generated the HS is taking for all noisy transient records. Then the overall features and the subtle features are extracted from

the noisy transients and the corresponded TFED, respectively. All the generated features are smoothed by means of a 4th-order, one-dimensional median filter. This yields six groups, as shown in block diagram (Fig. 6.3), of data subject to classifiers implementation.

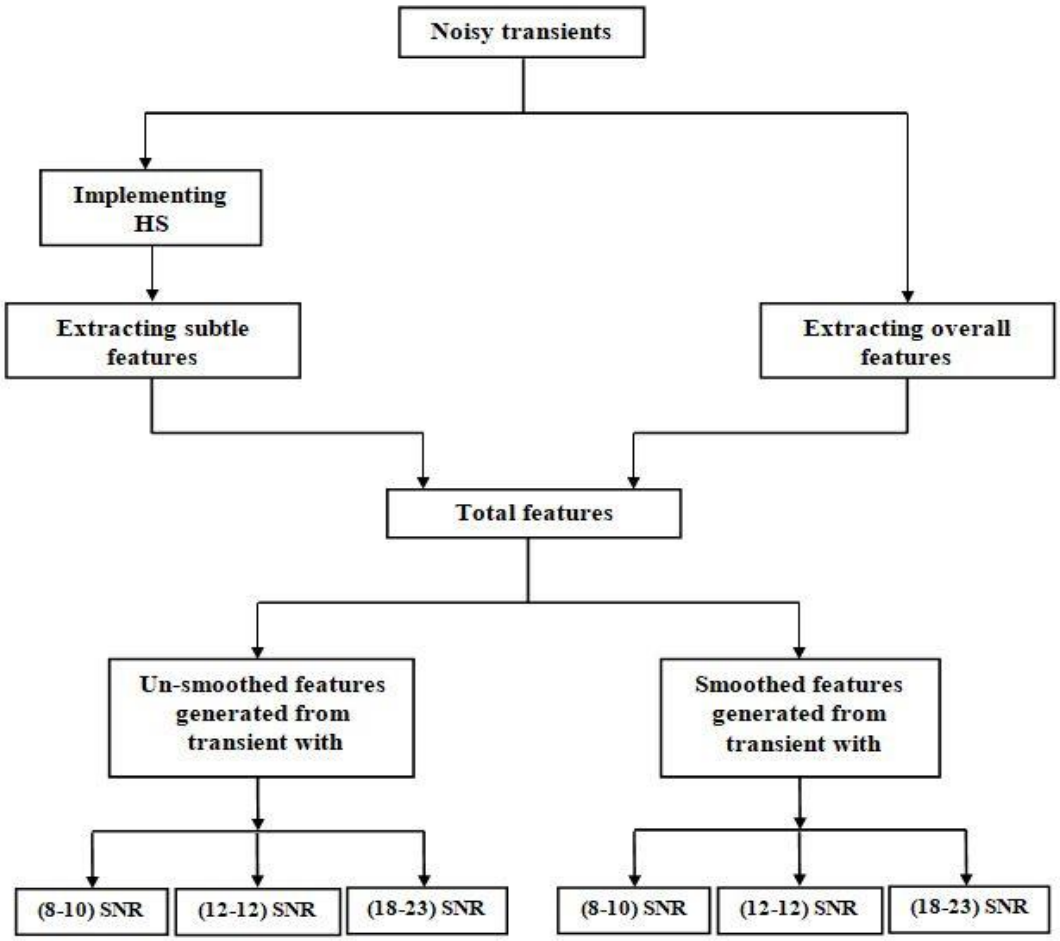


Figure 6.3 block diagram of process of adding noise to transient

The introduced three classifiers are applied to the six groups of the noisy transient features. Eighteen classification processes are executed to evaluate the effects of noisy signals on the classifiers' performances.

6.1 Unsmoothed features

6.1.1 Classifiers training for unsmoothed features

As it was done in the previous chapter, the classification process starts with dividing the data into two groups namely train data and tested data with (40%) and (60%) of the features' matrices data, respectively. The considered classifiers are trained by applying the learning processes to the 40% of the data. In this stage of classification process, forecasting criteria are generated to predict the performance of the classification. These criteria include train confusion matrix, ROC, AUC, and Scatter plot. As an example, Fig. 6.4 through Fig. 6.6 show the train confusion matrices that generating from training models of LDA, L-SVM, and Complex Tree classifiers, respectively. 40 % of 12-15 SNR un-smoothed transients are used in the training processes to produce the models.

The average accuracy of the model that is associated with the train confusion matrix, in Fig. 6.4, is 78.4%. In which the maximum true class records in the prediction are 59 records out of 60 records with 98.3% accuracy. However, the minimum true class records are 32 records with 53.3% accuracy. Class_1 (Huawei G5-A) and Class_8 (I-Phone 6S Plus-B) have the maximum and minimum true class records, respectively.

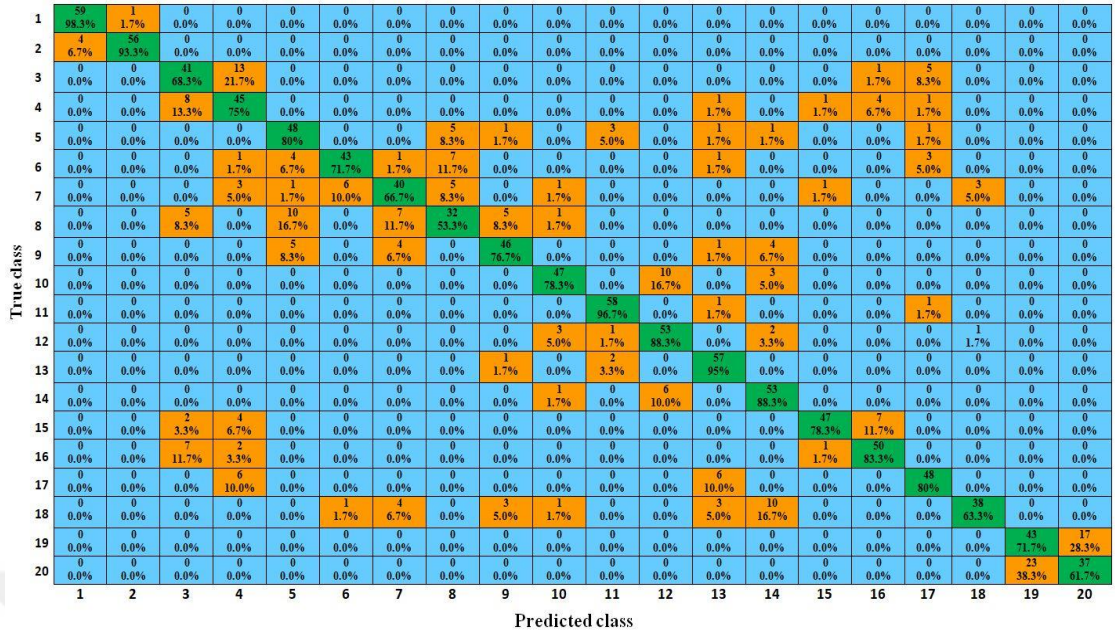


Figure 6.4 Train confusion matrix for 12-15 SNR un-smoothed transients via LDA.

The L-SVM classifier has the highest training accuracy (81.1%), in average, among the introduced classifiers as shown in Fig. 6.5. Sixty records of Class_1 are correctly predicted with 100% accuracy. However, 36 true class records which represent the minimum prediction accuracy (60%) belong to Class_8.

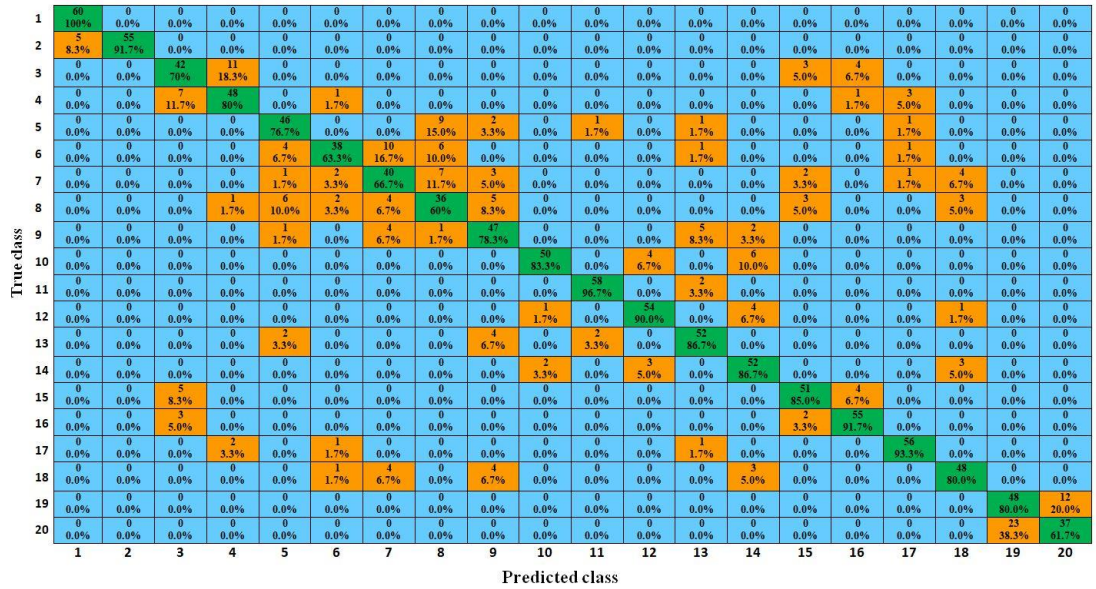


Figure 6.5 Train confusion matrix for 12-15 SNR un-smoothed transients via L-SVM.

The lowest training accuracy (73.3%) model, in average, is obtained by applying the Complex Tree classifier. In this model, the maximum training accuracy (95.0%) is associated with Class_1 (Huawei G5-A) and Class_2 (Huawei G5-B) whereas the minimum training accuracy (43.3%) is associated with Class_8.

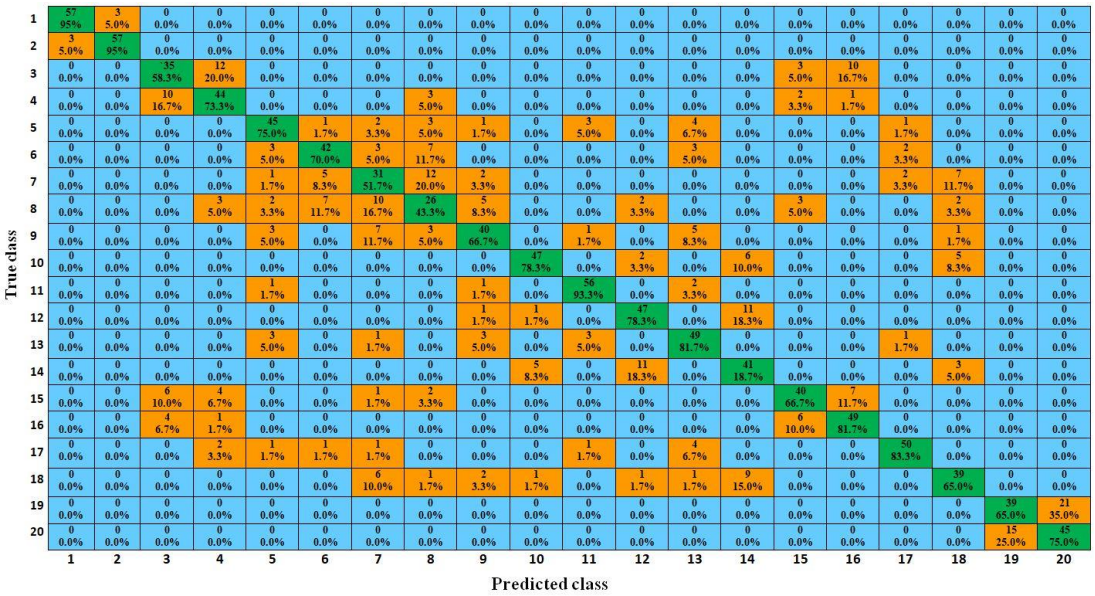


Figure 6.6 Train confusion matrix for 12-15 SNR un-smoothed transients via Complex Tree.

It's inferred from the train confusion matrices that Class_1 and Class_8 have the maximum and minimum true prediction accuracy, respectively, for all applied three classifiers with slightly different accuracy values. This means that the classifier training process depends on the robustness degree of the extracted features. Table 6.1 demonstrates the training accuracies of the three applied classifiers for unsmoothed transient features at different levels of SNR.

Table 6.1 Training accuracies of classifiers for unsmoothed features at different SNR levels.

Classifier	SNR level			
	8-10 dB	12-15 dB	18-23 dB	High SNR
L-SVM	73.5%	81.1%	87.8%	98.93%
C-Tree	61.9%	73.3%	79.7%	92.0%
LDA	71.0%	78.4%	82.8%	97.3%

It's obvious from the Table 6.1 that as level of SNR increases the training accuracy is increased. Moreover, it can be seen that the highest accuracy is associated with L-SVM in all SNR levels.

The ROC and AUC are other criteria by which the test classification performance can be predicted [48],[49]. The ROCs and AUCs of classifiers training process are illustrated in Fig. 6.7. It's demonstrated in the Figure that the AUC corresponding to L-SVM classifier is greater than the AUCs of other classifiers. This indicates better test classification performance of L-SVM.

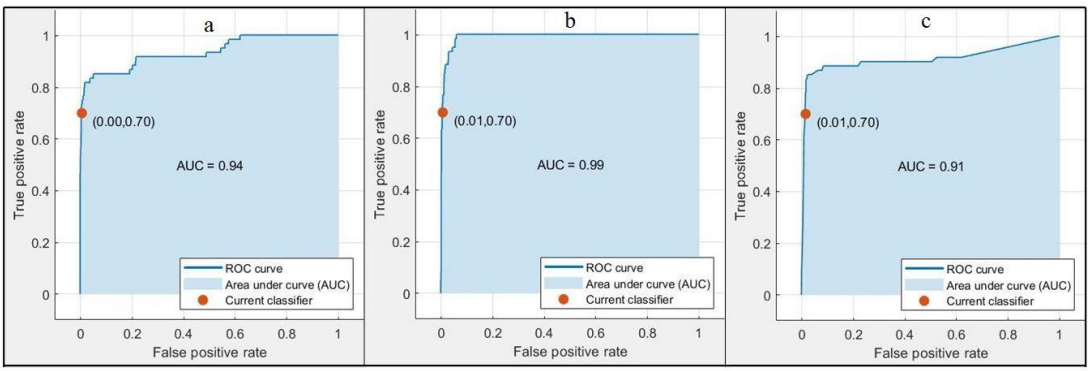


Figure 6.7 The ROC and AUC for 12-15 SNR un-smoothed arbitrary transients generated by a- LDA classifier, b- L-SVM classifier, and c- Complex Tree.

AUCs of classifiers training process for different SNR levels are illustrated in Table 6.2. The demonstrated AUCs are obtained, as an example, for Class_6 (I-Phone 6S-B). The AUCs values generally increased as SNR level increases for all applied

classifiers. At a high level of SNR, the AUC values reaches one when L-SVM is applied so that, in this case, the highest test performance can be predicted.

Table 6.2. Areas under ROC curves with different levels of SNR for different classifiers based on unsmoothed features.

Classifier	SNR			
	8-10	12-15	18-23	High SNR
L-SVM	0.98	0.99	0.99	1
C-Tree	0.89	0.91	0.94	0.99
LDA	0.94	0.94	0.95	0.99

By analyzing the training process predicting criteria (train confusion matrix accuracy, ROC and AUC) the test performances of the classifiers can be predicted. Based on the classifiers training accuracies (see Table 6.1) the test performances are expected to be slightly less than the corresponding train accuracies. Moreover, the highest expected classifier test performance will be associated with applying L-SVM classifier to high SNR level data. This can be deduced from the highest accuracy (98.93) in Table 6.1, and the highest AUC (1) in Table 6.2.

6.1.2 Classifier testing for unsmoothed features

The testing process of classification is accomplished by applying the introduced classifiers (LDA, L-SVM, and C-Tree) to the noisy test data based on generated training models. The noisy test data is chosen to be 60% of the total noisy data, such that 90 records of 150 total records are selected to be a noisy test data for each class.

The training models are generated for the three classifiers at three different levels of SNR, so nine different training models are generated in the noisy performance classification process. For each classifier, each training model is applied to the corresponding noisy test data to evaluate the noisy test performances.

The noisy test confusion matrices are shown in Fig. 6.8 through Fig. 6.10. The noisy test confusion matrix that is generated by applying the training model of LAD to 12-15 SNR un-smoothed noisy test data is shown in Fig. 6.8. In which the average test performance is 72.5.

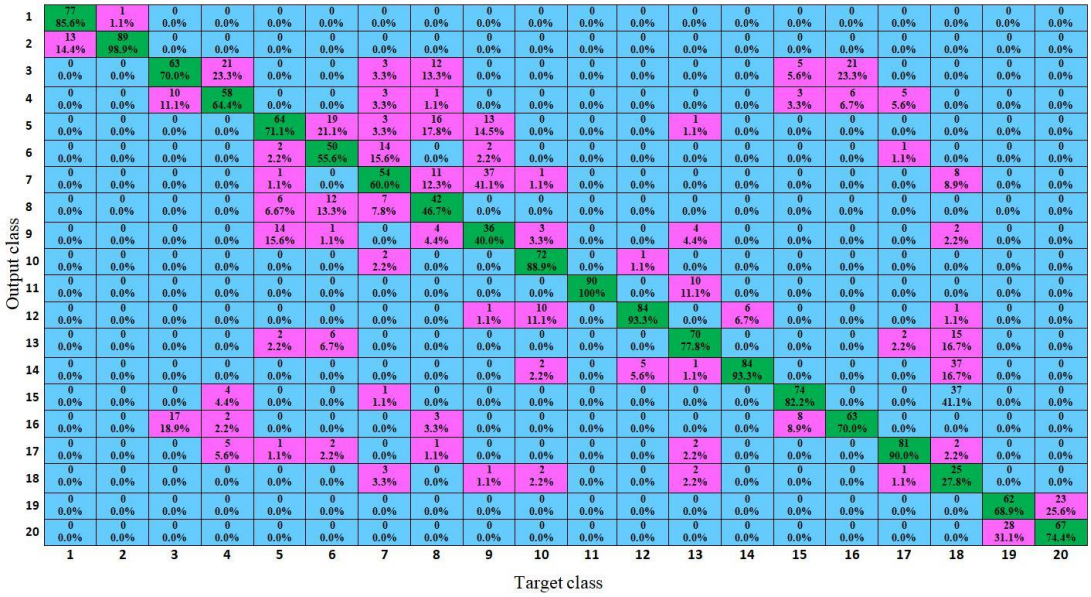


Figure 6.8 Test confusion matrix for 12-15 SNR un-smoothed transients via LDA.

Figure 6.9 show the noisy test confusion matrix with an average test performance (79.1%). This confusion matrix is generated by applying the training model of L-SVM to 12-15 SNR un-smoothed noisy test data. As it was predicted in the classifiers training process, the noisy test performance of the L-SVM (79.1%) is the highest test performance among the noisy test performances of the implemented classifiers. That is because training accuracy (81.1 %) and AUC (0.99) of L-SVM are the highest values of predicting criteria in the noisy data training process, as demonstrated in Table 6.1, and Table 6.2.

Output class \ Target class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	81 90%	2 2.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
2	9 10%	88 97.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
3	0 0.0%	0 0.0%	67 74.4%	16 17.8%	0 0.0%	0 0.0%	0 0.0%	4 4.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	5 5.6%	16 17.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
4	0 0.0%	0 0.0%	9 10.0%	71 78.9%	0 0.0%	0 0.0%	0 0.0%	2 2.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	4 4.4%	0 0.0%	1 1.1%	0 0.0%	0 0.0%	0 0.0%
5	0 0.0%	0 0.0%	0 0.0%	0 0.0%	71 78.9%	12 13.3%	3 3.3%	7 7.8%	5 5.6%	0 0.0%	1 1.1%	0 0.0%	4 4.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
6	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 1.1%	57 63.3%	5 5.6%	2 2.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	4 4.4%	0 0.0%	0 0.0%	0 0.0%
7	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 2.2%	70 77.8%	5 5.6%	39 43.3%	1 1.1%	0 0.0%	0 0.0%	1 1.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	6 6.7%	0 0.0%	0 0.0%
8	0 0.0%	0 0.0%	0 0.0%	0 0.0%	6 6.67%	12 13.3%	5 5.6%	53 58.9%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
9	0 0.0%	0 0.0%	0 0.0%	0 0.0%	6 6.67%	0 0.0%	1 1.1%	5 5.6%	41 45.6%	0 0.0%	0 0.0%	0 0.0%	7 7.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	12 13.3%	0 0.0%	0 0.0%
10	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 1.1%	1 1.1%	0 0.0%	80 88.9%	0 0.0%	1 1.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
11	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	89 98.9%	0 0.0%	6 6.7%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
12	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 1.1%	0 0.0%	85 94.5%	0 0.0%	1 1.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
13	0 0.0%	0 0.0%	0 0.0%	0 0.0%	6 6.67%	1 1.1%	0 0.0%	0 0.0%	2 2.2%	0 0.0%	1 1.1%	0 0.0%	71 78.9%	0 0.0%	0 0.0%	0 0.0%	2 2.2%	0 0.0%	0 0.0%	0 0.0%
14	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 2.2%	0 0.0%	4 4.4%	0 0.0%	86 95.6%	0 0.0%	0 0.0%	0 0.0%	15 16.7%	0 0.0%	0 0.0%
15	0 0.0%	0 0.0%	0 0.0%	2 2.2%	0 0.0%	0 0.0%	4 4.4%	7 7.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	74 82.2%	7 7.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
16	0 0.0%	0 0.0%	14 15.6%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 1.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	7 7.8%	67 74.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
17	0 0.0%	0 0.0%	0 0.0%	1 1.1%	0 0.0%	6 6.7%	0 0.0%	1 1.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 1.1%	0 0.0%	0 0.0%	0 0.0%	83 92.2%	0 0.0%	0 0.0%	0 0.0%
18	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 2.2%	2 2.2%	3 3.3%	6 6.7%	0 0.0%	0 0.0%	0 0.0%	3 3.3%	0 0.0%	0 0.0%	0 0.0%	57 63.3%	0 0.0%	0 0.0%
19	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	69 76.7%	27 30.0%
20	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	21 23.3%	63 70.0%

Figure 6.9 Test confusion matrix for 12-15 SNR un-smoothed transients via L-SVM.

Figure 6.10 illustrates the noisy test confusion matrix generated by applying the training model of C-Tree to 12-15 SNR un-smoothed noisy test data. The average of C-Tree classifier test performance, in this case, is 66.2 %, which is the lowest classifier performance among the applied classifier in the noise test performance.

Output class \ Target class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	85 94.4%	3 3.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
2	5 5.6%	87 96.7%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
3	0 0.0%	0 0.0%	69 76.7%	18 20.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 3.3%	19 21.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
4	0 0.0%	0 0.0%	6 6.67%	61 67.8%	0 0.0%	0 0.0%	0 0.0%	7 7.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	11 12.3%	2 2.2%	1 1.1%	0 0.0%	0 0.0%	0 0.0%
5	0 0.0%	0 0.0%	0 0.0%	0 0.0%	58 64.4%	11 12.3%	0 0.0%	8 8.9%	4 4.4%	0 0.0%	16 17.8%	0 0.0%	19 21.1%	5 5.6%	0 0.0%	0 0.0%	2 2.2%	2 2.2%	0 0.0%	0 0.0%
6	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 3.3%	50 55.6%	17 18.9%	7 7.8%	2 2.2%	0 0.0%	0 0.0%	0 0.0%	3 3.3%	0 0.0%	0 0.0%	0 0.0%	4 4.4%	0 0.0%	0 0.0%	0 0.0%
7	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	44 48.9%	7 7.8%	35 38.9%	6 6.67%	0 0.0%	0 0.0%	1 1.1%	0 0.0%	19 21.1%	0 0.0%	0 0.0%	3 3.3%	0 0.0%	0 0.0%
8	0 0.0%	0 0.0%	0 0.0%	0 0.0%	7 7.8%	10 11.1%	23 25.6%	43 47.8%	3 3.3%	3 3.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 1.1%	0 0.0%	3 3.3%	0 0.0%	0 0.0%	0 0.0%
9	0 0.0%	0 0.0%	0 0.0%	0 0.0%	10 11.1%	0 0.0%	0 0.0%	3 3.3%	32 35.6%	0 0.0%	6 6.67%	0 0.0%	7 7.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	13 14.4%	0 0.0%	0 0.0%
10	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	48 53.3%	0 0.0%	2 2.2%	0 0.0%	3 3.3%	0 0.0%	0 0.0%	0 0.0%	2 2.2%	0 0.0%	0 0.0%
11	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	66 73.3%	0 0.0%	18 20.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
12	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	9 10.0%	0 0.0%	72 80.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	5 5.6%	0 0.0%	0 0.0%
13	0 0.0%	0 0.0%	0 0.0%	0 0.0%	9 10.0%	5 5.6%	1 1.1%	1 1.1%	2 2.2%	0 0.0%	2 2.2%	0 0.0%	41 45.6%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
14	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	5 5.6%	0 0.0%	16 17.8%	0 0.0%	80 88.9%	0 0.0%	0 0.0%	0 0.0%	9 10.0%	0 0.0%	0 0.0%
15	0 0.0%	0 0.0%	2 2.2%	8 8.9%	0 0.0%	0 0.0%	0 0.0%	7 7.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	52 57.8%	14 15.6%	2 2.2%	0 0.0%	0 0.0%	0 0.0%
16	0 0.0%	0 0.0%	13 14.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 3.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	4 4.4%	55 61.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
17	0 0.0%	0 0.0%	0 0.0%	3 3.3%	3 3.3%	14 15.6%	0 0.0%	4 4.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 1.1%	0 0.0%	0 0.0%	0 0.0%	78 86.7%	0 0.0%	0 0.0%	0 0.0%
18	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	5 5.6%	3 3.3%	12 13.3%	19 21.1%	0 0.0%	0 0.0%	0 0.0%	2 2.2%	0 0.0%	0 0.0%	0 0.0%	56 62.2%	0 0.0%	0 0.0%
19	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	53 58.9%	28 31.1%
20	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	37 41.1%	62 68.9%

Figure 6.10 Test confusion matrix for 12-15 SNR un-smoothed transients via C-Tree.

Considering un-smoothed transient features and based on different levels of SNR, the introduced classifiers test performances are illustrated in Table 6.3. It's deduced from the Table that the classifiers' performances are increased as the SNR level increases. However, there are different effects on the classifiers' performances due to the level of SNR. For instant, The L-SVM classifier has the highest test performance at low SNR levels, followed by LDA classifier, and then C-Tree which has the lowest performance, whereas at high level of SNR, as illustrated in the Table, the C-Tree has the highest test performance. This is due to the capabilities of the classifiers at different level of data interference.

Table 6.3. Classifiers' test performances with different levels of SNR (un-smoothed features).

Classifier	SNR level			
	8-10 dB	12-15 dB	18-23 dB	High SNR
L-SVM	72.5%	79.1%	85.3%	97.2%
C-Tree	57.1%	66.2%	80.9%	97.59%
LDA	68.8%	72.5%	72.8%	91.3%

6.2 Classifier testing for smoothed features

In order to decrease the level of interference degree among the features of noisy transient signals and to investigate the smoothing process on the classifiers' performances, the same order (4th-order) of one-dimensional median filter that is applied at high SNR level data is implemented to noisy data. After applying the median filter to the different levels of noisy features, smoothed features are generated. The considered classifiers are trained based on 40%, (60 records out of 150 records), of the smoothed noisy data and then tested by the rest data. The test confusion matrices of the three introduced classifiers (LDA, L-SVM, and C-Tree) applied to 12-15 SNR smoothed data are demonstrated in Fig. 6.10 through Fig. 6.12.

The LDA classifier test performances for each class are shown in the matrix diagonal in Fig. 6.11, where the average LDA test performance for all classes is 77.8%. It's notice, comparing to the average test performance of LDA for 12-15 SNR unsmoothed is (72.5%), that the smoothing process improves the applied classifier performance.

Output class \ Target class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	89 98.9%	1 1.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	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%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
3	0 0.0%	0 0.0%	68 75.6%	25 27.8%	0 0.0%	0 0.0%	0 0.0%	2 2.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	9 10.0%	16 17.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
4	0 0.0%	0 0.0%	5 5.6%	48 53.3%	0 0.0%	0 0.0%	0 0.0%	2 2.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 2.2%	0 0.0%	5 5.6%	0 0.0%	0 0.0%	0 0.0%
5	0 0.0%	0 0.0%	0 0.0%	0 0.0%	76 84.4%	1 1.1%	1 1.1%	8 8.9%	18 20.0%	0 0.0%	1 1.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
6	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	56 45.6%	13 14.4%	1 1.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
7	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	9 10.0%	63 70.0%	1 1.1%	30 33.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	8 8.9%	0 0.0%	0 0.0%
8	0 0.0%	0 0.0%	0 0.0%	0 0.0%	6 6.7%	22 22.4%	8 8.9%	76 84.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
9	0 0.0%	0 0.0%	0 0.0%	0 0.0%	5 5.6%	1 1.1%	0 0.0%	0 0.0%	42 46.7%	18 20.0%	0 0.0%	0 0.0%	1 1.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
10	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	72 80.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 1.1%	0 0.0%	0 0.0%
11	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	90 100%	0 0.0%	20 22.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
12	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	88 97.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	7 7.8%	0 0.0%	0 0.0%
13	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 3.3%	1 1.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	69 76.7%	0 0.0%	0 0.0%	0 0.0%	1 1.1%	17 18.9%	0 0.0%	0 0.0%
14	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 2.2%	0 0.0%	90 100%	0 0.0%	0 0.0%	0 0.0%	35 38.9%	0 0.0%	0 0.0%
15	0 0.0%	0 0.0%	0 0.0%	5 5.6%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	76 84.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
16	0 0.0%	0 0.0%	17 18.9%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 3.3%	74 82.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%
17	0 0.0%	0 0.0%	0 0.0%	12 13.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	84 93.3%	0 0.0%	0 0.0%	0 0.0%
18	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	5 5.6%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	22 24.4%	0 0.0%	0 0.0%
19	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	63 70.0%	24 26,7%
20	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	27 30.0%	66 73.3%

Figure 6.11 Test confusion matrix for 12-15 SNR smoothed transients via LDA.

Test confusion matrix that is obtained by applying L-SVM classifier to the 12-15 SNR smoothed data is shown in Fig. 6.12. In which the average classifier test performance is 82.2 %. The confusion matrix demonstrates an improving in the test performance result comparing to L_SVM test performance for unsmoothed data that shown in Fig. 6.9.

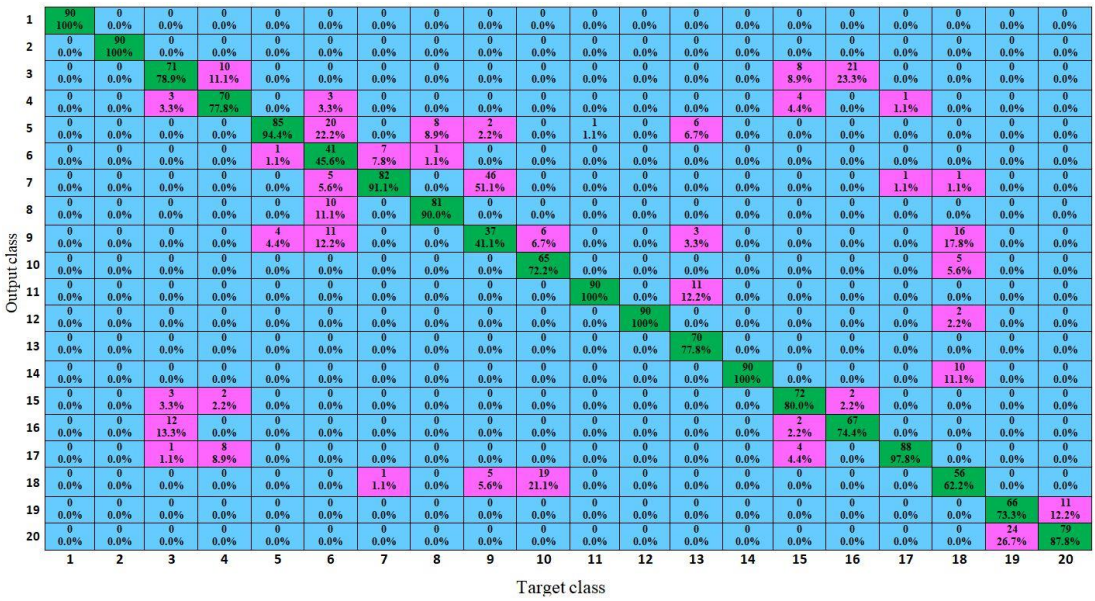


Figure 6.12 Test confusion matrix for 12-15 SNR smoothed transients via L-SVM.

Figure 6.13 shows the confusion matrix for 12-15 SNR smoothed transients obtained via applying C-Tree. The average C-Tree test performance for 12-15 SNR smoothed noisy data is 68.8%. This performance, at this level of SNR for smoothed data, is the lowest classifier performance among the test performances of the three implemented classifiers. However, the average C-Tree classifier test performance is improved comparing to 66.2 % test performance of the same SNR level that shown in Fig. 6.10.

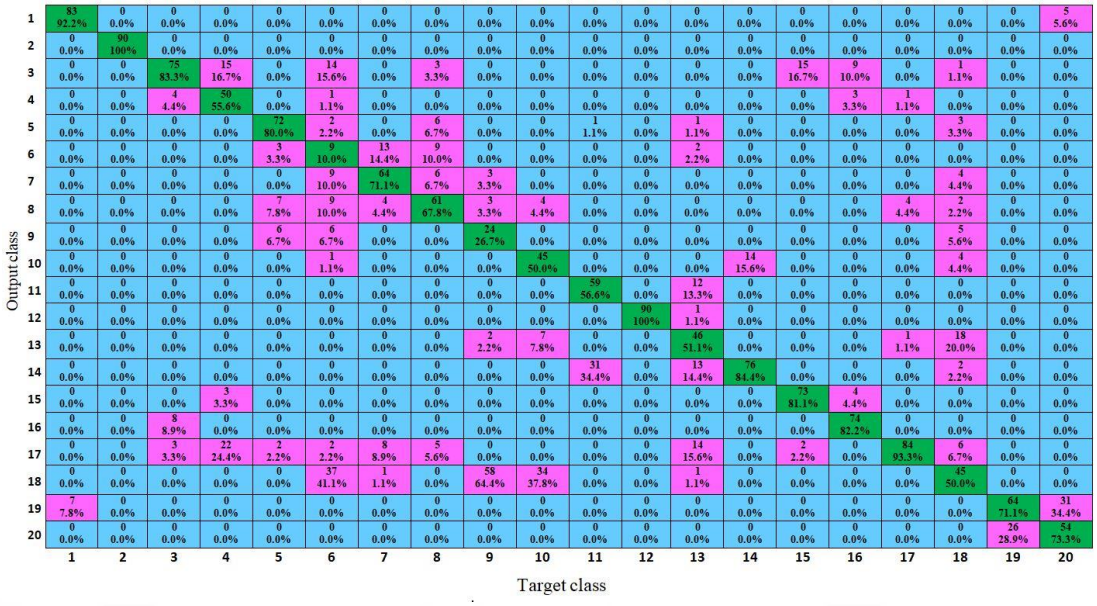


Figure 6.13 Test confusion matrix for 12-15 SNR smoothed transients via C-Tree.

The classifiers test performances that are obtained from applying the classifiers to smoothed noisy transient features at different SNR levels are shown in Table 6.4. Generally, as inferred from the Table, the classifiers test performances is increased, as the SNR level increases. However, they become close to each other at high SNR level. As shown in the table, The L-SVM has the highest test performance at low SNR levels, whereas the test performance of the C-Tree is rapidly improved as SNR level increases. This is due to the variation of the classifiers effectiveness against the noisy data.

Table 6.4. Classifiers' test performances with different levels of SNR (smoothed features).

Classifier	SNR level			
	8-10 dB	12-15 dB	18-23 dB	High SNR
L-SVM	79.3%	82.1%	90.53%	99.8%
C-Tree	66.8%	68.8%	85.4%	99.6%
LDA	76.6%	77.8%	83.6%	99.7%

It’s concluded from thesmoothing process that the process generally improves the classifiers test performances as shown in Fig. 6.7 through Fig. 6.13. However, there are some classification of certain classes are negatively affected by applying the smoothing process. For instant, as a worst case, the test performance of C-Tree of class-6 (I-Phone 6s-B) is decreased from 50% to 9% after applying the smoothed process as shown in Fig. 6.10 and Fig. 6.13, respectively. This misclassification which might be considered as a disadvantage of using feature smoothing results from convergence of feature data of each class to its median. This convergence might lead to more interference with other classes’ features.

CHAPTER 7

CONCLUSION

In this work, extensive data from BT devices are considered in RF fingerprinting implementation. BT signals are collected from different cell phones' brands, models and serial numbers. Two devices of ten different cell phones, each with one hundred and fifty records, are considered in this study. The collected BT transmission signals are de-noised by BPF. The transient signals of the de-noised transmission signals are detected by using a novel transient detection technique based on energy envelope. By using this technique, the start and end points of the transient are detected. However, it is necessary to improve the proposed techniques for both more practical implementations and lower computational loads. The algorithm is based on the use of the local energy envelope maxima. The multiple local energy maximums leads to interference in the transient durations of different classes. Concequently, the transient duration cannot be used as a robust feature. This is resolved by exploiting the transient features to re-capture the transient end. Therefore, the energy envelope based transient detection algorithm has been improved by introducing several more stages. The results indicate that the technique is efficient and applicable for all records of the huge data.

Based on the EMD method, the HHT is applied to the detected transient signals to obtain their TFED. The EMD applied to the detected transient signal of each BT signal in order to generate its IMFs. The Hilbert transform is needed to be applied to each extracted IMFs to generate their instantaneous characteristics. Having obtained the characteristics of each IMF, the Hilbert spectrum (HS) is generated.

Thirteen features are extracted from three feature groups namely, the transient signal and its energy envelope, summation of TFED of the transient signal along time axis,

and summation of TFED of the transient signal along frequency axis. The extracted features are pre-processed to evaluate their usability. Among thirteen features, the STD of transient instantaneous phase, the entropy of the transient instantaneous phase, the duration of the transient, and variance of summation of energy along time axis are considered as the most robust features. These features are dominant in constructing the signal RF fingerprint for each record.

Before the classification, the robustness of the extracted features are investigated by using the Box plot demonstration, scatter plot, or conventional plot. By means of using the box plot a clear picture of the interference among classes feature can be seen. In addition, the percentage of the data concentration can be recognized. By using the conventional plot, the interfered records and the number of interfered records can be known.

Smoothing process is needed to recognize the interference among classes feature and enhance the robustness of the features. As a part of smoothing we zoomed the extracted features. We can clearly recognise the interference among classes feature and see the difference after and before zooming.

In order to identify the twenty considered classes, different version of five classifiers are applied to the extracted features. Each extracted feature data of each class is divided into two data groups namely, training set and testing set. First we introduced 40% and 60% of the total data, train and test set respectively for smoothed and unsmoothed features. And then we used 20% and 80% for train and test data respectively for unsmoothed features. The purpose of splitting data set into training and test data is to see the realistic performance of the classifiers. We introduced the different percent of train data and test data for the smoothed and unsmoothed features to investigate the effect of the data division size. The average accuracy of classifiers' based on 40%-train 60%-test unsmoothed RF fingerprint is 93.07%, whereas, the average accuracy of classifiers' accuracies based on 20%-train 80%-test unsmoothed RF fingerprint is 92.32%. It's concluded from Table 5.4 and Table 5.5 that the test accuracies are slightly increased as the train data percent increases.

The classification results demonstrate that the method is effective in BT device identification, and it achieved quite high performance. After a comparison between

the classification accuracies of classifiers based on unsmoothed and smoothed features, it's concluded that the classifiers' accuracies are increased after applying the smoothing process. Based on smoothed features,the average accuracies of the ten applied are 99.85% as train accuracy and 99.72% as test accuracy, whereas, they demonstrate, based on unsmoothed features, 93.19% as train accuracy and 93.07% as test accuracy.

The performances of the introduced classifers are evaluated at different SNR levels. Fetures of three levels of noisy transient signals are devided into smoothed and unsmoothed data. It's concluded from the performance evaluation that the classifiers test performances demonstrate high performances at high level of SNR comparing to low levels. It' seen from the noise performance results that the test performances of L-SVM are the highest performance values at all levels of SNR for both smoothed and unsmoothed data, whereas the test performance of the C-Tree is rapidly improved as SNR level increases. This is due to the variation of the classifiers effectiveness against the noisy data. It can be also concluded from the introducing of the smoothing process that the process generally improves the classifiers test performances. However, there are some classification of certain classes are negatively affected by applying the smoothing process.

REFERENCES

[1] Y. Shiu., and S.Chang. “Physical Layer Security In Wireless Networks: A Tutorial’’ IEEE Wireless Communications, vol. 18(2), pp. 66-74, April 2011.

[2] O. Ureten and N. Serinken “Wireless Security Through RF Fingerprinting,” Can. J. Elect. Comput. Eng., Vol. 32(1), pp. 27-33, Winter 2007.

[3] B. Danev. et al. "Attacks on physical-layer identification." Proceedings of the third ACM conference on Wireless network security. ACM, pp. 89-98, 2010.

[4] C. Sperandio and P. Flikkema, “Wireless Physical-Layer Security via Transmit Precoding Over Dispersive Channels: Optimum Linear Eavesdropping,” MILCOM 2002. Proceedings. Vol. 2. IEEE, 2002.

[5] J. L. Padilla, et al. "RF fingerprint measurements for the identification of devices in wireless communication networks based on feature reduction and subspace transformation." Measurement, vol. 58, pp. 468-475, 2014.

[6] B.W. Ramsey, et al. “Wireless infrastructure protection using low-cost radio frequency fingerprinting receivers. ” international journal of critical infrastructure protection, vol. 8 pp. 27-39, 2015.

[7] W. Albazrqaoe, “A Study of Bluetooth Frequency Hopping Sequence: Modeling and a Practical Attack,”. Michigan State University. Computer Science, 2011.

[8] Introduction to bluetooth Device testing. National Instruments. http://download.ni.com/evaluation/rf/intro_to_bluetooth_test.pdf http://www.ni.com/rfacademy/.

[9] J.C Haartsen. “The Bluetooth radio system.” IEEE personal communications, vol. 7(1), pp.28-36, 2000.

[10] K. Leung, “Bluetooth Personal Area Network: An Investigation of new networking technology, ” in McMaster University. Available: Accessed: Oct. 17, 2015.

[11] J. MANDER, D. PICOPOULOS, “Bluetooth piconet applications. London’s Global University, London, 2005.

[12] 12.4 Bluetooth Protocol. https://flylib.com/books/en/4.215.1.116/1/

[13] J. Hall, M. Barbeau, and E. Kranakis, “Detection of Transient in Radio Frequency Fingerprinting Using Signal Phase.” Proceedings of ASTED International Conference on Wireless and Optical Communications (WOC), Banff, Alberta, pp. 13-18, 2003.

[14] O. Ureten and N. Serinken, “Bayesian detection of radio transmitter turn-on transients.”, *Proc. NSIP99*, pp. 830-834, 1999.

[15] O. Ureten and N. Serinken, “Detection of radio transmitter turn-on transients” Electronics Letters, volume 35(23), pp. 1996-1997, 1999.

[16] Y. Yuan, Z. Huang, H. Wu, X. Wang. “Specific Emitter Identification Based on Hilbert–Huang Transform-Based Time–Frequency–Energy Distribution Features,” IET Communications, Vol. 8, Iss. 13, pp. 2404–2412, 2014.

[17] SU. Rehman, K. Sowerby, and C. Coghill. “RF Fingerprint Extraction from the Energy Envelope of an Instantaneous Transient Signal”. Department of Electrical and Computer Engineering University of Auckland, New Zealand, Australian Communications Theory Workshop (AusCTW), IEEE, pp. 90-95. 2012.

[18] Y.J. Yuan, Z. Huang, and Z.C.Sha. "Specific emitter identification based on transient energy trajectory." Progress In Electromagnetics Research, vol. 44 , pp. 67-82, 2013.

[19] N. E. Huang, *“*Introduction to the Hilbert Huang Transform and Its Related Mathematical Problems,” In Hilbert–Huang transform and its applications. Pp. 1-26, 2014.

[20] MD. K. Islam Molla, and K. Hirose, *“*Hilbert Spectrum in Time-Frequency Representation of Audio Signals Considering Disjoint Orthogonality,” advances in adaptive data analysis vol. 2, no. 3, pp. 313–336, 2010.

[21] N. E. Huang, X. Chen, M. T. Lo, and Z. Wu. “On Hilbert Spectral Representation a True Time- Frequency Representation for Nonlinear and Non-Stationary Data,” Advances in Adaptive Data Analysis Vol. 3, Nos. 1 & 2, 63–93, (2011).

[22] N. E. Huang, *et al.* (1998). “The Empirical Mode Decomposition and Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis”. Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences. Vol. 454. No. 1971. pp. 903-995. The Royal Society, 1998.

[23] R. Korkikian, D. Naccache, and G. Almeida, “Instantaneous frequency analysis. ” Data Communication Networking (DCNET), International Conference on. IEEE pp. IS-11, 2013.

[24] D. Kim and H-Seok Oh, “EMD: A Package for Empirical Mode Decomposition and Hilbert Spectrum,” The R Journal Vol. 1/1, pp. 40-46, May 2009.

[25] G. Wang, et al. "On intrinsic mode function." Advances in Adaptive Data Analysis, 2(03), pp. 277-293, 2010.

[26] Md. K. Islam Molla, A. Sumi and M. S. Rahman, “ Analysis of Temperature Change under Global Warming Impact using Empirical Mode Decomposition,” International Journal of Information Technology Vol. 3(2), pp 131-139, 2006 .

[27] Md. K. Islam Molla, K. Hirose, and N. Minematsu, “Localization Based Separation of Mixed Audio Signals with Binary Masking of Hilbert Spectrum,” 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings. Vol. 5. pp. V-V. IEEE, 2006.

[28] C. S. Turner, “Hilbert Transforms, Analytic Functions, and Analytic Signals,” Retrieved from personal. atl. bellsouth. net/p/h/physics/hilberttransforms. pdf (2005).

[29] Md. I. Khan, Md. E. Hamid, and T. Nakai. “Systolic Phase Detection from Pulsed Doppler Ultrasound Signal Using EMD DHT Based Approach,” International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.7, No.5 pp.207-216, 2014.

[30] J.P. Ángel-López, and N. A. de la Peńa. "Kinematic Hand Analysis Using Motion Capture Technology. ” In VI Latin American Congress on Biomedical

Engineering CLAIB 2014, Paraná, Argentina 29, 30 & 31 October 2014, pp. 257-260. Springer, Cham, 2015.

[31] P. Singh, S. D. Joshi, R. K. Patney , and K. Saha, " The Fourier Decomposition Method for Nonlinear and Non-Stationary Time Series Analysis," Proc. R. Soc. A, vol. 473(2199), 20160871. 2017.

[32] M.I.A. Khan, M.E. Hamid, and T. Naka, "EEMD and DHT Based Solution to Detect Systolic Phase from Doppler Ultrasound Signal," International Conference on Materials, Electronics & Information Engineering, ICMEIE-2015 05-06 June, 2015,University of Rajshahi, Bangladesh.

[33] N. E. Huang, and Z. Wu, "A Review on Hilbert Huang Transform Method and Its Applications to Geophysical Studies," Reviews of geophysics 46(2), 2008.

[34] Md. E. Hamid, Md. K. Islam Molla, Md. I. Khan, *T. Nakai*," Speech Enhancement Using Hilbert Spectrum and Wavelet Packet Based Soft-Thresholding," Science Journal of Circuits, Systems and Signal Processing 4(1), pp. 1-8. 2015.

[35] R. Jalaja, B. Biswal, " Power Quality Event Classification Using Hilbert Huang Transform," Springer Berlin Heidelberg, Proceedings of the International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA), Vol. 199, pp 153-161, 2013.

[36] B. Friedlander, B. Porat, "Detection of Transient Signals by the Gabor Representation," IEEE transactions on acoustics, speech, and signal processing 37.2 (1989) ,pp. 169-180.

[37] L. U. Iurinic, R. G. Ferraz, A. S. Bretas, A. L. Bettiol, L. F. N. dos Passos, A. Carniato, I. K. Khairalla, R. Z. Homma, and A. Baader, "Fault Induced Transient Detection in Power Systems: An Adaptive Approach Considering Noisy Environment," IEEE Transactions on Power Delivery, DOI 10.1109/TPWRD.2016.2589265.

[38] G. Ernst, " Heart Rate Variability," London: Springer, 2014.

[39] B. Qian, and K. Rasheed, "Hurst Exponent and Financial Market Predictability ". In Proceedings of The 2nd IASTED international conference on financial engineering and applications, pp. 203-209. 2004.

[40] F. Cervantes-De la Torre, J. I. Gonzalez-Trejo, C. A. Real-Ramrez, and L. F. Hoyos-Reyes, ” Fractal dimension algorithms and their application to time series associated with natural phenomena” 4th National Meeting in Chaos, Complex System and Time Series, Journal of Physics: Conference Series 475(1), p. 012002, (2013).

[41] G. G. Coyt, and et al, “Higuchi’s Method applied to the Detection of Periodic Components in Time Series and Its Application to Seismograms” COMPLEX SYSTEMS Revista Mexicana de Fısica S 59 (1), pp. 1–6, 2013.

[42] K.J. Ellis and N. Serinken, “Characteristics of radio transmitter fingerprints,” Radio Science, Vol. 36(4), pp. 585-59, 2001.

[43] Ali, A. M., Uzundurukan, E., Kara, A.: “Improvements on transient signal detection for RF fingerprinting*”*. In Signal Processing and Communications Applications Conference, Antalya, Turkey, pp. 1–4, May 2017.

[44] E. Uzundurukan, A . M. Ali, and A . Kara,: “Design of low-costmodular RF front end for RF fingerprinting of Bluetooth signals*”*. In Signal Processing and and communications applications conference, Antalya, Turkey, May 2017, 15, pp. 1–4, IEEE.

[45] J. L. Hintze, “Chapter 152 Box plots number cruncher statistical systems*”* Hintze, Jerry L. "User’s Guide V." (2007).

[46] M. A. Friedl, C. E. Brodley, “Decision tree classification of land cover from remotely sensed data. *”*. Remote sensing of environment 61.3 (1997): 399-409.

[47] S. Wang, B. Meng, and H. Tian. "A modeling method based on Wavelet Support Vector Machine." Control and Decision Conference (CCDC), pp. 3113-31162010 Chinese. IEEE, 2010.

[48] C. Cortes, and M. Mohri, "AUC optimization vs. error rate minimization." Advances in neural information processing systems,pp. 313–320, 2004.

[49] T. Fawcett, "An introduction to ROC analysis." Pattern recognition letters, vol. 27, no. 8,pp. 861–874, 2006.