# IEEE Signal Processing Cup 2016
# Team - MGLS

Nilushika Udayangani, Nimeshika Udayangani, Sahani Goonetilleke, Piyumika Bandula, Amalika Lahiruni, Nipuni Mello and Kapila Chandika Wavegedara

*Department of Electronic and Telecommunication Engineering*
*University of Moratuwa,*
*Sri Lanka*

***Abstract* - Extracting the Electric Network Frequency (ENF) fluctuations from an audio recording and comparing it to a reference database is a new approach in performing forensic digital audio authentication. The problem statement of the IEEE SP cup 2016 competition relates to time-varying location-dependent signature of power grids as it becomes intrinsically captured in media recordings, due to direct or indirect influences from the respective power grid. In this project signal processing and information security/forensics are collectively elaborated.**

**The objective of this project is to implement and design a range of programs and algorithms for capturing and extracting ENF signals for a novel ENF based application. Without relying on the availability of concurrent power references, an ENF pattern extracted from a given digital recording was examined to infer the power grid in which the recording was done. The statistical features of ENF pattern variations for this region-of recording-identification process were exploited. With these features a multiclass SVM classifier, which is able to identify the grid of recording of a given signal, was implemented. The latter part of the project involves with circuit design and data analysis for ENF acquisition.**

**A Graphical User Interface (GUI) was developed to display the most probable three, grid-of-origins for a given recording. The whole process was tested and evaluated for the test data given.**

***Index Terms – Signal processing algorithms, MATLAB, Hardware***

## I. Introduction

Electric Network Frequency (ENF) can be used as a forensic tool with conjunction with a database of ground truth ENF measurements. A common approach, uses the ENF signal, to geo-locate a digital recording, exploiting the fact that ENF variations associated with a power grid is almost identical in all locations of the same grid and show some sort of unique variations for a given grid, due to inherent load variations and controlled mechanisms within the grid.

This project examines such approach to geolocate a multimedia recording using its embedded ENF, in terms of their grid-of-origin. Without concurrent power references, this method identifies the grid in which the ENF containing signal is recorded.

We were given power and audio recordings from nine different grids around the world. There were significantly different in the nature and manner of the ENF variations in different grids. Due to this reason, statistical features extracted from the grids are also different. Hence processing the ENF signal from the given recording, to extract these statistical features; and matching with those of feasible grids, paves way to the identification of the grid in which the recording was obtained, thus geo-locating the recoding. After this we implemented a machine learning system that learns the differences between those statistical features from different grids and used it to classify the ENF signals in terms of their region-of-recording.

Due to significant differences of “clean” ENF data extracted from power recordings and “noisy” ENF data extracted from audio recordings, we developed two separate classifier systems for audio recordings and power recordings. The system makes a prior analysis to identify the recording in terms of power or audio; and then passes through the respective classifier to obtain the results.

The rest of this paper is organized as follow:

Section II describes the extraction of ENF signals from both training and testing data sets and significant differences in the nature and the manner ENF variations among different grids are analysed. Section III describes the location classification system we have implemented while Section IV presents the circuit design and data analysis for ENF acquisition and finally Section V gives the conclusion of the paper.

## II. Extraction of ENF Signals

In our preliminary research, we came across many feasible methods for ENF pattern extraction from a digital recording. They included spectrum combining method [1] and frequency demodulation method [2]. Of them the method presented by Cooper [3] was adapted due to its low complexity.

*A. Approaches Used to Extract ENF Signals*.

All the audio and power recordings from the provided training set were initially sliced into segments of length of 600,000 samples. We chose this length since all the given test recordings were of that length. (Sampling rate of 1000Hz corresponds to a time duration of 10 minutes per segment of each recording) The signal decimation step was foregone since the 1000Hz is permitted to approximate frequencies of 50/60 Hz as it does not result in a longer computational time.

*1. Bandpass Filtering*

In some recordings of both 50Hz and 60Hz grids and some power recordings, the majority of ENF information were not present at the fundamental frequencies (i.e. at 50Hz and 60Hz respectively), but at higher harmonics. For example, we observed that the $10^{th}$ and $31^{st}$ power recordings from the practice data set had the ENF information clearly saturated near the third harmonic (150Hz). Therefore prior to determining filter specifications such as the cut-off frequency and the passband for the bandpass filter, we analysed the signal as follows.

First, the FFT of each 10 minute long segments of the recording was calculated and plotted. Then it was determined if the recording was an audio or power recording. If it was a power recording, the dominant frequency of 50 or 60 Hz was identified. Usually the audio recording produced a more scattered and wriggly FFT plot, while that of the power recording indicated one or two clearly identifiable peaks at its fundamental frequency or second or third harmonics. (For audio recordings this differentiation is not required as it is explicitly given which is audio and power)

Then each power recording of both 50Hz and 60Hz were bandpass filtered with the cut-off frequencies set to 49.5-50.5Hz and 59.5-60.5Hz respectively; while that of the audio recording was 49.95Hz and 50.05Hz, a comparatively narrower passband.

*2. Dividing into Partially Overlapping frames*

The filtered signal was then divided into M number of frames of length L*D each, which partially overlapped eachother. The hop size, L, was set to the sampling rate of 1000Hz. Hence the final ENF pattern resolution obtained was 1 second. I.e. finally we obtained 600 ENF values for a 10 minute long recording. The factor D, which determines the total length of the frame, L*D; was set to 3.

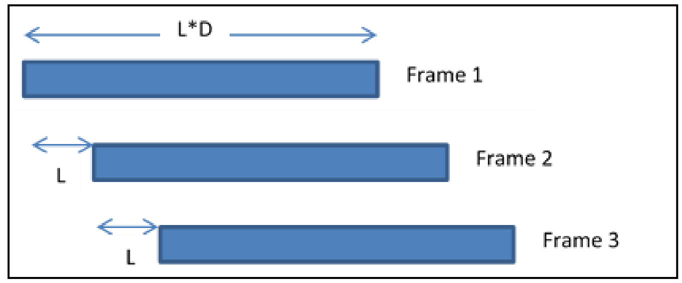


*Figure 1: Overlapping frames*

The overlapping between adjacent frames govern the smoothness of the final ENF pattern. For a smooth ENF extraction from audio recordings D should be large. When examining outputs for different values of D, the value of 3 was selected since the length overlapping two adjacent frames was $2f_s$ in this case ($f_s$ = 1000Hz sampling rate).

*3. Windowing Each Frame*

Windowing each frame with appropriate window function and zero padding factors was the next step. The best window function for estimation of maximum amplitude frequency via Quadratic Interpolation is Blackman window with zero padding factor close to 1.3. Blackman window with zero padding factor of 4*600,000 was used for the power recordings. It was observed that better results were given when rectangular window function with zero padding factor set to 49*600,000 was used for audio recordings.

*4. DFT for Each Windowed Frame*

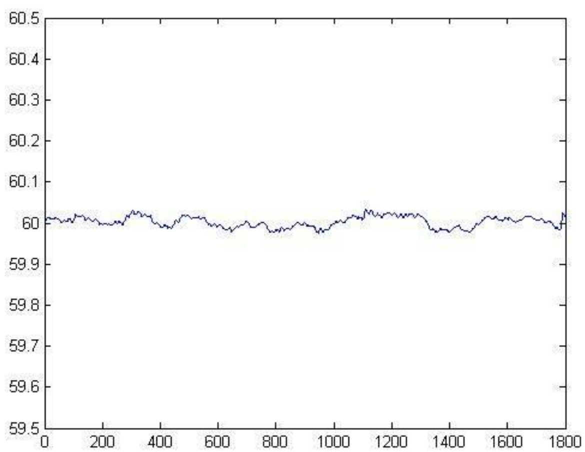
DFT for each windowed frame was obtained by performing FFT. Since zero padding factors mentioned above were used, the number of points used to perform FFT for power recordings and audio recording were 5,000 length (frame) and 50,000 length (frame) respectively.

*5. Peak Frequency Estimation via Quadratic Interpolation*

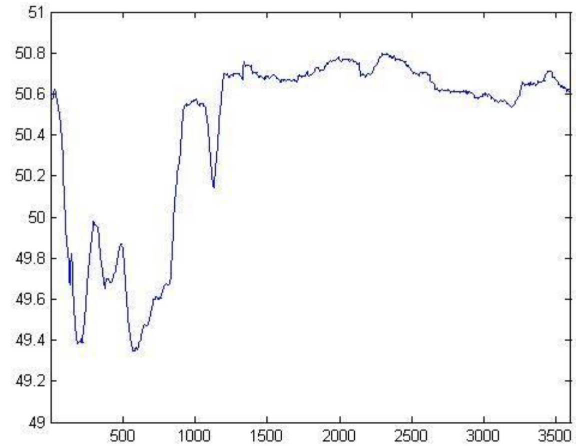
For each frequency spectrum a bin number with maximum amplitude was found. As the first step, the frequency associated with that bin number was calculated and it was considered as the peak frequency. 60 or 50 values were obtained for all the frames irrespective of small variations. It was observed that it is unlikely that the peak frequency coincided exactly with a DFT frequency bin, because moderate zero padding factors and sampling frequency of 1000Hz was used. Hence the resolution of FFT was proved to be poor. Two solutions were identified for this issue. They are making the FFT resolution high by using much higher zero padding factors and small sampling frequency or making estimations using numerical methods like Quadratic Interpolation and spline estimation etc. The latter was adopted here.

First, frequency bin N was located with maximum amplitude and amplitudes of frequency bins of N-1 and N+1 were taken into account to perform Quadratic Interpolation. Next the estimated peak frequency was stored as the ENF value for the corresponding frame so that an extracted ENF pattern of M ENF values was finally obtained. Here, M implies the number of frames.

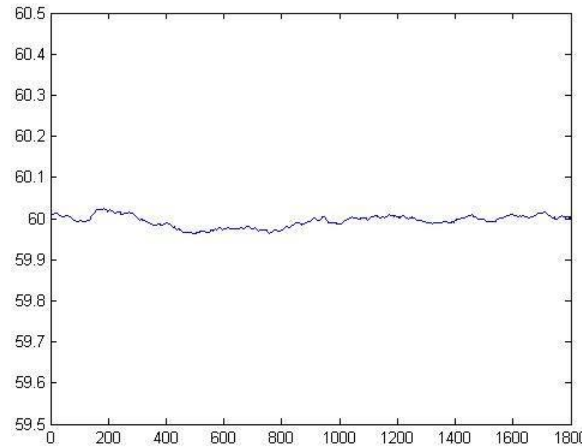
Clear ENF signals were given by this method for power recordings. ENF signals were corrupted by several spikes for audio recordings. Extracted ENF pattern for power recordings was passed through a medium filter of order 40. Spikes were replaced by the median of 40 close by ENF values. The spikes were not completely removed for small orders of the medium filter. It was observed that when the order of the median filter was increased, details of the ENF pattern were reduced. A moderate value for order was adopted by trial and error method (i.e. 40).



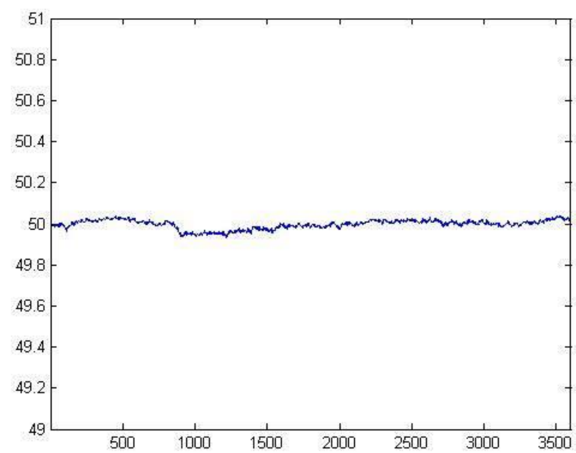

*A*



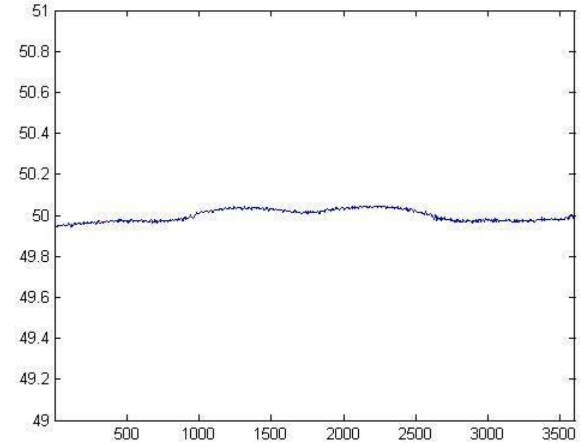

*B*



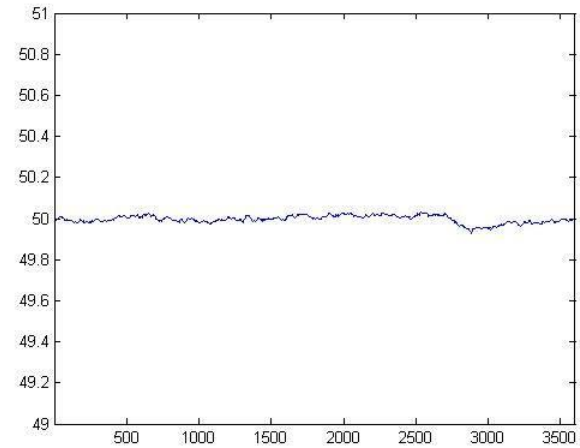

*C*



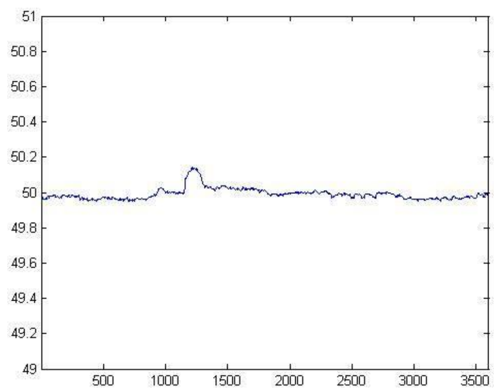

*D*



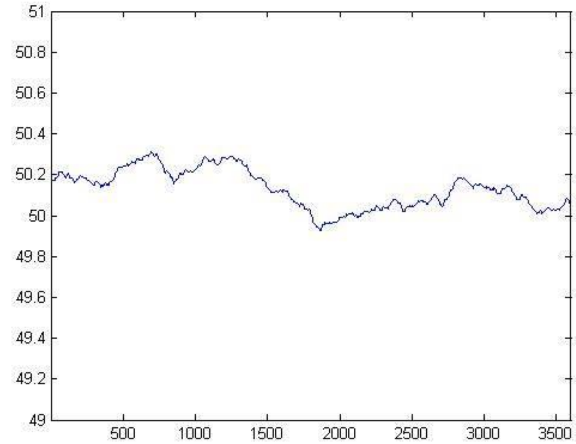

*E*



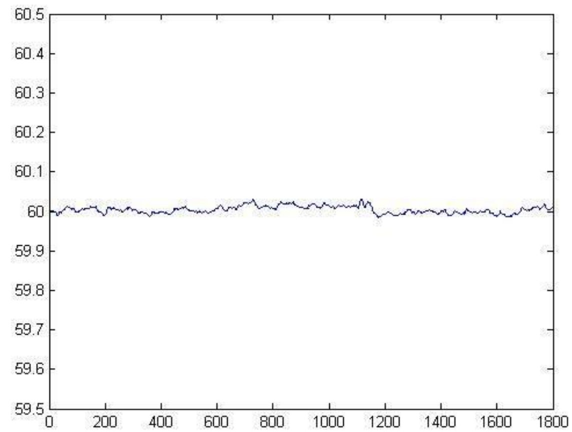

*F*




*G*




*H*




*I*

*Figure2: ENF Signals Extracted from the Recordings*

*B. Initial observations on extracted ENF patterns*:

From the ENF patterns of nine grids it can be easily identified A, C and I are of 60Hz grids while the others are of 50Hz grids. In figure 2 ENF pattern variations from A, C and I are plotted using 30 minutes long power recordings whereas other ENF patterns plotted using one hour long power recordings. Note that amplitude scale axis is also different for 60Hz grids and 50Hz grids.

60Hz grids seems to be more controlled than 50Hz grids, and also show high similarity in the manner of their variations. A has slightly more ripples compared with C and I. It can be seen B, D, E, F, G and H are of 50Hz grids. Among them ENF from D,E and F seems to be more controlled than other 50Hz grids, but not as much as in 60Hz grids. B and H shows comparatively more variations. B has outliers that drop around 1Hz, a characteristic that does not appear in other ENF signal samples. ENF signal mean of H grid can be seen to be above 50Hz most of the time.

## III. LOCATION CLASSIFICATION/IDENTIFICATION SYSTEM

### *A. Design for the System*

The classification system used is based on multiclass SVM classifiers using statistical features extracted from ENF patterns. [5] The classification systems for audio and power were built separately in order to increase accuracy.

*Feature Extraction:*

From each ENF signal segment 13 quantitative features were extracted.

1) Mean of ENF segment
2) log(variance) of ENF segment
3) log(range) of ENF segment
4) log(variance) of approximation after L-level wavelet analysis (L = 9)

5-13) log(variance) of 9 levels of detail signals computed through L-level wavelet analysis from coarser to finer (L = 9)

By applying log operator to the dynamic range and variance features, their order of magnitude was clearer, and separability was enhanced. [5]

By observing ENF patterns extracted from power recordings of each grid, it was identified that the mean, variance and dynamic range were simple, yet powerful candidates for features. To understand, first we built our multiclass SVM classifier system using only the above three features. To train the system, from each grid, only 30 random power recordings of 10 minute duration were used from each training set. Thus from the remaining power recordings, the accuracy of the SVM classifier system was tested based on the same three features only. The accuracy given was 62% (Audio examples were not tested for this classifier system)

Since we needed to increase accuracy, and also obtain robust results for audio as well, we included the remaining above-mentioned4-13 features; extracting through 9-level wavelet decomposition.

In order to use training data to train the classifier, we took 30 power recordings of length 10 minutes, from each of the 9 grids, and extracted their features while storing them grid wise. Also from each grid, 6 audio recordings of length 10 minutes were done likewise.

The next step was to normalize these feature values to the range of [-100,100] by using a linear scale. [5] Since number of examples from both audio and power recordings were equal, we calculated the mean over all power and audio recording data separately. Then we found the maximum value for deviation of each feature value from the mean of the feature considered; and stored these mean deviation values of each feature in power data and audio data as two separate sets. Then we normalized each feature value using this information and was stored once more.

*SVM Multiclass Classifier:*

According to the discussion in paper [5], a classifier was trained using support vector machine (SVM). Multiclass classifier system was made for these nine classes after completion of training 36 numbers of binary classes. Each binary classifier is trained on one of the 36 possible pairs of classes.

Training multiclass classifier for power recording system:

Linear, polynomial, RBF, mlp and Quadratic kernel in MATLAB [6] was tried for each 36 possible pair of classes. Based on the accuracy from results linear kernel was adopted to separate A and C power grid pair. Rest of the 36 pairs was trained using RBF kernel with sigma value of 2.9.

Training multiclass classifier for audio recording system:

After testing each of the kernel functions to each of the possible binary classifier, polynomial kernel with order 2 was adopted for A and I grids. Linear kernel was used for all other 35 pairs.

Testing with multiclass classifier system:

When a 10 minute long recording was given from training data set, it is identified as audio or power prior to ENF extraction. Then 13 features are extracted and normalized using corresponding means and mean deviations.

If the recording is identified as an audio recording, it is tested with each binary classifier and give a one mark for winning class from each binary classification task. If one class emerges as winner for all 8 binary classification tasks which it includes, then it gets a total of 8 marks. Ultimately all the classes were given a mark between 0 and 8. Class claiming the maximum mark was eligible for the winning class. If it is a power recording carried out, a similar marking procedure using multiclass classifier is trained for power recording system.

*B. Classification Results Obtained on the Testing Data*

Practice data results and accuracy:

Among the given 50 recordings, we identified 28 recordings were clear power recordings. After ENF extraction by the proposed method, test examples were separately given to two multiclass classifier systems. I.e. audio or power based on if testing example is the respective recording.

The accuracy obtained was 60% and the result is as follows

Practice Results:

AHCFF,GBCBD,AFGDC,IGBAE,DBBFD,HEFGB,DDEGG,EDBHI,HGECF,FBGEB

Accuracy: 60%

Test-data results:

The provided test data set was separated on the basis of power or audio recording and then classified separately. Then finally combined two results. The corresponding result is as follows.

Test results

BDDBD,FGDAF,FGGBE,BFCEH,DHHDG,FFGAI,DBFGE,IGCBD,EGGBE,EGEAG,GIGGG,HAEFC,BFFDG,CECFI,EICGF,BDBDF,DFDFG,EABAH,FHDBA,GBFBG

*C. Essential Information about our Submitted Software Codes*

In the submitted software code, there are 2 folders, namely, "System_for_Practice_dataset" and "System_for_Testing_dataset". Each contains entire set of codes which help in classifying the data set implied by its name. (I.e. power or audio)

When navigating in the practice data set folder, there are 2 subfolders "Audio_SVM_for_Practice_audio recordings" and "Power_SVM_for_practice_power recordings". When in the "Audio_SVM_for_Practice_audio recordings" folder, open "Audio_classifier.m" MATLAB file. Once run it will write results for only audio recordings at corresponding index (1-50) into a "Practice_result.mat" file. Simply it fills zeros for indexes corresponding to power recordings, while one of the letters A to I are written at indexes corresponding to audio recordings. Then insert this mat file into "Power_SVM_for_Practice_power recordings" folder and run "Power_classifier.m". Then at places kept with zeros, this code will write results for power recordings into "Practice_result.mat" file and also in the MATLAB command window.

Follow same steps to obtain results for testing dataset, while navigating through "System_for_Test_dataset" folder. "GUI" folder is included a user friendly GUI so that any single recording can be classified. When run 'text1.mat' file and select recording from drop down menu, it will display first 3 mostly likely grid labels.

## IV. Circuit Design and Data Analysis for ENF Acquisition

The transformer was used to step down the Voltage from the local Voltage of 230V to a safe Voltage. In the process of authenticating an audio recording other than the construction of the reference database based on the given ENF signals, capturing an ENF signal throughout the course of recording the audio file and classifying it also plays a major role. The part 2 of the open competition built on the Part 1 serves above fact bringing out the synergy of sensing, processing, and learning.

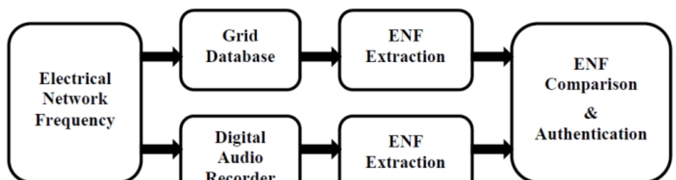



*Figure 3: Hardware Approach Taken in Capturing ENF in Sound Recording.*

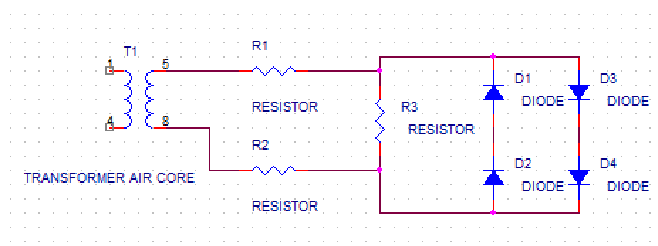



*Figure 4: Schematic Diagram*

## V. Conclusion

The accuracy of 60% was obtained for the practice results. Training the multiclass classifier systems separately for power and audio recordings, contributed to increase the overall accuracy. Since ENF pattern extracted from power recording is very clean comparatively with it from audio recordings, it gives more accurate results for region-of-recording identification of power recordings than audio recordings. The experience gained as a team while working on this challenge is invaluable.

## Acknowledgment

We would like to thank our supervisor, friends and relatives for their support and encouragement to complete this challenge.

## References

[1] A. Hajj-Ahmad, R. Garg and Min Wu, "Spectrum Combining for ENF Signal Estimation", IEEE Signal Processing Letters, vol. 20, no. 9, pp. 885-888, 2013.

[2] L. Dosiek, "Extracting Electrical Network Frequency From Digital Recordings Using Frequency Demodulation", IEEE Signal Processing Letters, vol. 22, no. 6, pp. 691-695, 2015.

[3] A. Cooper, "An automated approach to the Electric Network Frequency (ENF) criterion - Theory and practice", IJSLL, vol. 16, no. 2, 2010.

[4] Y. Liu, Z. Yuan, P. Markham, R. Conners and Y. Liu, "Application of Power System Frequency for Digital Audio Authentication", *IEEE Transactions on Power Delivery*, vol. 27, no. 4, pp. 1820-1828, 2012.

[5] A. Hajj-Ahmad, R. Garg and Min Wu, "ENF-Based Region-of-Recording Identification for Media Signals", IEEE Trans.Inform.Forensic Secur., vol. 10, no. 6, pp. 1125-1136, 2015.

[6] P. Top, M. Bell, E. Coyle and O. Wasynczuk, "Observing the Power Grid: Working Toward a More Intelligent, Efficient, and Reliable Smart Grid with Increasing User Visibility", IEEE Signal Process. Mag., vol. 29, no. 5, pp. 24-32, 2012.

[7] 2016. [Online]. Available: https://www.ruor.uottawa.ca/bitstream/10393/24383/3/El_Gemayel_Tarek_2013_thesis.pdf. [Accessed: 14- Jan- 2016].

[8] Wikipedia, "Electrical network frequency analysis", 2016. [Online]. Available: https://en.wikipedia.org/wiki/Electrical_network_frequency_analysis. [Accessed: 14- Jan- 2016].

[9] Trace.tennessee.edu, 2016. [Online]. Available: http://trace.tennessee.edu/cgi/viewcontent.cgi?article=2871&context=utk_gradthes. [Accessed: 14- Jan- 2016].

[10] 2016. [Online]. Available: http://www.luju.ro/static/files/2012/februarie_2012/12/teza_masterat.pdf. [Accessed: 14- Jan- 2016].

[11] 2016. [Online]. Available: http://www.emo.org.tr/ekler/909fcedd9fd41ab_ek.pdf. [Accessed: 14- Jan- 2016].

[12] 2016. [Online]. Available: http://www.emo.org.tr/ekler/909fcedd9fd41ab_ek.pdf. [Accessed: 14- Jan- 2016].

[13] collector, "NFI enf collector", SourceForge, 2013. [Online]. Available: http://sourceforge.net/projects/nfienfcollector/?source=directory. [Accessed: 14- Jan- 2016].

[14] Code.google.com, "/ - digital-audio-forensic - digital audio authentication - Google Project Hosting", 2016. [Online]. Available: https://code.google.com/p/digital-audio-forensic/source/browse/#svn%2Ftrunk%2Fmatlab%20project%2FENFExtractor. [Accessed: 14- Jan- 2016].

[15] Mathworks.com, "How to extract ENF (Electrical Network Frequency) form audio signal. I want extracted ENF pattern from the audio signal using... - MATLAB Answers - MATLAB Central", 2016. [Online]. Available: http://www.mathworks.com/matlabcentral/answers/142645-how-to-extract-enf-electrical-network-frequency-form-audio-signal-i-want-extracted-enf-pattern-fr. [Accessed: 14- Jan- 2016].

[16] A. Cooper, "The Electric Network Frequency (ENF) as an Aid to Authenticating Forensic Digital Audio Recordings – an Automated Approach", Audio Engineering Society, 2008.

[17] Mast.umd.edu, "Electric Network Frequency (ENF) Overview", 2016. [Online]. Available: https://www.mast.umd.edu/index.php/enf-menu. [Accessed: 14- Jan- 2016].

[18] A. Hajj-Ahmad, R. Garg and Min Wu, "ENF-Based Region-of-Recording Identification for Media Signals", IEEE Trans.Inform.Forensic Secur., vol. 10, no. 6, pp. 1125-1136, 2015.

[19] D. Bykhovsky and A. Cohen, "Electrical Network Frequency (ENF) Maximum-Likelihood Estimation Via a Multitone Harmonic Model", IEEE Trans.Inform.Forensic Secur., vol. 8, no. 5, pp. 744-753, 2013.

[20] A. Hajj-Ahmad, R. Garg and Min Wu, "ENF-Based Region-of-Recording Identification for Media Signals", IEEE Trans.Inform.Forensic Secur., vol. 10, no. 6, pp. 1125-1136, 2015.

[21] M. Stamm, Min Wu and K. Liu, "Information Forensics: An Overview of the First Decade", IEEE Access, vol. 1, pp. 167-200, 2013

**H. D. Nimeshika Udayangani** is an undergraduate (level 2) of the Department of Electronic and Telecommunication Engineering, University of Moratuwa, Sri Lanka. She is a student member of the IEEE student chapter of University of Moratuwa. She is currently engaged in a number of projects based on Digital Signal Processing and Embedded Electronic Systems. She has participated in IEEEXtreme 9.0, IEEE Electronic Design Competition 2015 and Sri Lanka Robotics Challenge 2015. Her primary research interests include new technologies for modern signal processing and electronic industry. She is exceptionally talented in painting and drawing.

**H. D. Nilushika Udayangani** is an undergraduate (level 2) of the Department of Electronic and Telecommunication Engineering, University of Moratuwa, Sri Lanka. She is a student member of the IEEE student chapter of University of Moratuwa. She is currently engaged in a number of projects based on Digital Signal Processing and Embedded Electronic Systems. She has participated in IEEEXtreme 9.0, IEEE Electronic Design Competition 2015 and Sri Lanka Robotics Challenge 2015. Her primary research interests include new technologies for modern signal processing and electronic industry. She is exceptionally talented in painting and drawing. H. D. Nimeshika Udayangani and H. D. Nilushika Udayangani are twin sisters.

**R. Sahani L. D. S. Goonetilleke** is an undergraduate (level 2) of the Department of Electronic and Telecommunication Engineering, University of Moratuwa, Sri Lanka. She is a member of the IEEE and a student member of the IEEE WIE student chapter of University of Moratuwa. Her previous projects include QRS detection, Sound Sensitive Lamp, Black Grain detector and dynamic target tracking robot development. She is currently engaged in a project on power amplifiers. She has participated in IEEEXtreme 8.0, IEEEXtreme 9.0, IEEE Electronic Design Competition 2015 and Sri Lanka Robotics Challenge 2015. Her research interests include telecommunication, image and signal processing, algorithms and embedded electronics.

**S. Piyumika T. Bandula** is an undergraduate (level 2) of the Department of Electronic and Telecommunication Engineering, Bio Medical Specialization, University of Moratuwa, Sri Lanka. She is a student member of the IEEE student chapter of University of Moratuwa. She has engaged in the projects; QRS detection algorithm, Cost effective checkers board, Cerebral Palsy curing toy production, RGB colour analyser that are based on Electronics and Bio Medical Engineering. Her primary research interests include Bio Medical image processing, Tissue Engineering and Algorithms and software development.

**W. A. K. Amalika Lahiruni** is an undergraduate (Level 2) of the Department of Electronic and Telecommunication Engineering, Bio-Medical Specialization, University of Moratuwa, Sri Lanka. She is a student member of the IEEE student chapter of University of Moratuwa. She is currently engaged in a number of projects; QRS detection algorithm, wireless racing car robot, Cerebral Palsy curing toy production and a mini project based on Power amplifiers. Her primary research interests include Bio Medical image processing and signal processing, new technologies for modern signal processing, Algorithms and software development

**G. D. C. Nipuni B. Mello** is an undergraduate (level 2) of the Department of Electronic and Telecommunication Engineering, University of Moratuwa, Sri Lanka. She is a student member of the IEEE, WIE student chapter of University of Moratuwa. She is currently engaged in a number of projects based on Internet of Things, sign language translator for interaction with dumb people based on Digital Signal Processing and a mini project based on Power amplifiers. She has been working on projects such as FM Transmitter, RGB Colour Analyser for low cost and high accuracy and a dynamic target-tracking robot. Her primary research interests include new technologies for modern telecom industry and astrophysics.

**Kapila Chandika Wavegedara** received the B.Sc. degree in Electrical and Electronics Engineering from the University of Peradeniya, Sri Lanka and the M.Eng degree in Telecommunications from Asian Institute of Engineering (AIT), Thailand in 1998 and 2001, respectively. He obtained the Ph.D. degree in Electrical and Computer Engineering from the University of British Columbia (UBC), Vancouver, Canada in 2008. He has been a Senior Lecturer at the Dept. of Electronics and Telecommunication, University of Moratuwa, Sri Lanka since April 2008. His main research interests lie in the area of Wireless Communications (Physical Layer).