IEEE Signal Processing Cup 2016
Team Ravan

*University of Moratuwa, Sri Lanka

Dilin Dampahalage* Dulanga Weerakoon* Chathura Abeyrathne* Pankaja Pathegama* Nipun Mahesha*

Supuni Gunasekara* Dulara De Zoysa* Yohan Nugera* Kasun Maduranga* Dasun Madushanka*

Dr.Chandika Wavegedara*

***Abstract*—This report describes a method which is implemented for the IEEE signal Processing Cup 2016, that is to extract power signatures from a given media signal and identify where the signal was recorded in order to be used for forensic applications. The primary task of this project was to design a system which can identify the captured location of a given multimedia sample based on the Electrical network frequency signals (ENF) signal embedded in it. A special characteristic of ENF is its time varying nature which allows it to be used in such applications. At the preliminary stage, the ENF signals needs to be extracted in order to be further processed. In the method proposed, the extraction of the ENF signals undergoes two stages. First the FFT mean amplitudes at 50 Hz and its harmonics is compared with mean amplitude at 60 Hz and its harmonics in order to determine the nominal frequency. Once nominal frequency is determined as 50Hz or 60Hz, a spectrogram based spectral combining approach [1] is used for the ENF signal estimation. Following the extraction of the ENF signal, the signal should be classified. Multiclass Support Vector Machine (SVM) classifier is used for the above task. For the training Audio and Power Recordings of 9 known power grids were used. Furthermore the system was trained with some statistical features of these training data such as Nominal frequency, Frequency range of ENF signal and Variance in a logarithmic scale. The extracted ENF signals of the multimedia recordings were then entered to the trained system for grid classification. Then probability scores for each grid are taken from the system. A threshold was used to determine that the signal belongs to a given grid or None of the above. Highest overall accuracy of 72% was achieved for the given dataset when multiclass Support Vector Machine (SVM) was used as the classifier. The secondary task was to implement a hardware solution that could record ENF from a standard power outlet. The device which was built using the Arduino platform was successful in obtaining ENF from the domestic supply and providing it to terminal software in the computer.**

***Index Terms*—Signal Processing, Electric Network Frequency, Support Vector Machines**

I. Introduction

Electric Network Frequency (ENF), the supply frequency in a power distribution network [1] has been used as forensic analysis tool for over the past few decades. Since the fluctuation of the ENF signal differs from one power distributor to another, ENF signal can be used as a tool to identify the location where a particular signal has been recorded. Therefore this technology can be used as a powerful tool in forensic applications such as terrorist attacks, ransom demands etc.

First task of this project was to build a system which can extract the ENF signals from multimedia signals accurately and identify the location or power grid which it belonged to. For this entire process, accuracy and the reliability would be the main concern since these identification systems will be used in forensic applications.

Second task of the project was to build a hardware device which can record power signals from a standard power outlet. Recorded power signals were further analyzed and results were compared with the features of the given signals of nine grids. This report presents the design details for both software based grid identification system and hardware device which is used to record power signals. Designed grid identification system is presented as a MATLAB based Graphical User Interface (GUI). For the first task, the system has two main parts, train the system from given known signals using a machine learning algorithm and identify a given signal from and unknown power grid using the trained system. Each of this part contains three main steps for a given signal,

- Identifying the nominal frequency (50Hz or 60Hz).
- Extract ENF Signal
- Training /Classification

Machine learning system was trained using some statistical features of extracted ENF signals. That trained model was then used for the identification of a given signal.

A description of the ENF extracting algorithm and machine learning algorithms, their corresponding accuracies are stated in the report.

II. Extraction of ENF Signal

It was identified that the ENF signal appears around the nominal frequency(50/60Hz) and its harmonic bands [1]. Furthermore it was observed that the ENF signal strengths

University of Moratuwa, Sri Lanka

around different harmonics vary, and this is due to differ recording environments and devices used. The implemen ENF extraction procedure undergoes 2 steps. Those are,

1) Nominal Frequency Detection
2) ENF extraction using Spectral Combining approach

Analyzing Fast Fourier Transform (FFT) and Spectrograms Power Signal Recordings and an Audio Signal Recordings was clearly identified that in Power Signals ENF is mai observed in its nominal frequency band and in Audio Sign ENF can also be observed in its harmonic bands.

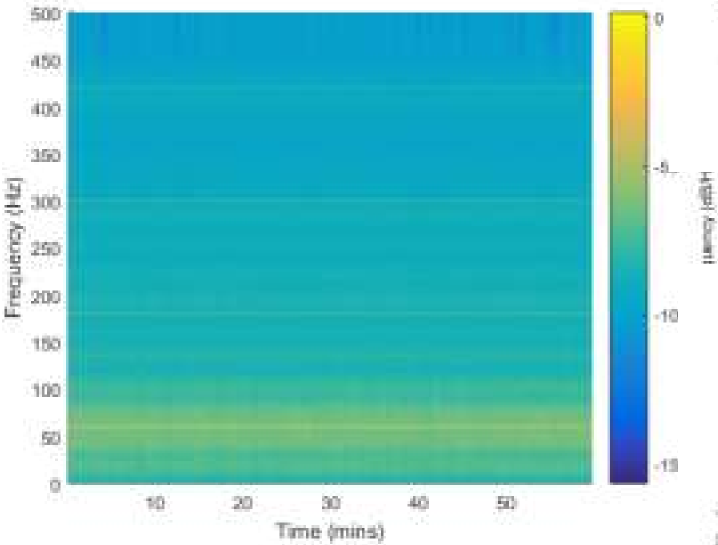

Fig. 1. Spectrogram of power signal Train_Grid_A_P1.wav

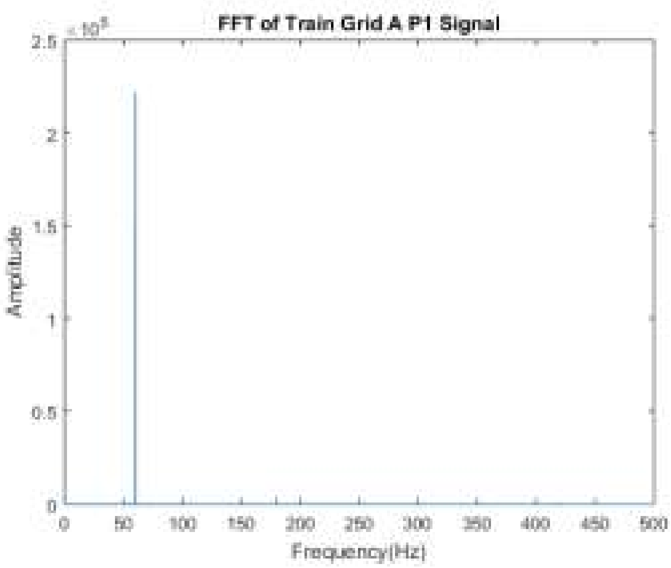

Fig. 2. FFT of power signal Train_Grid_A_P1.wav

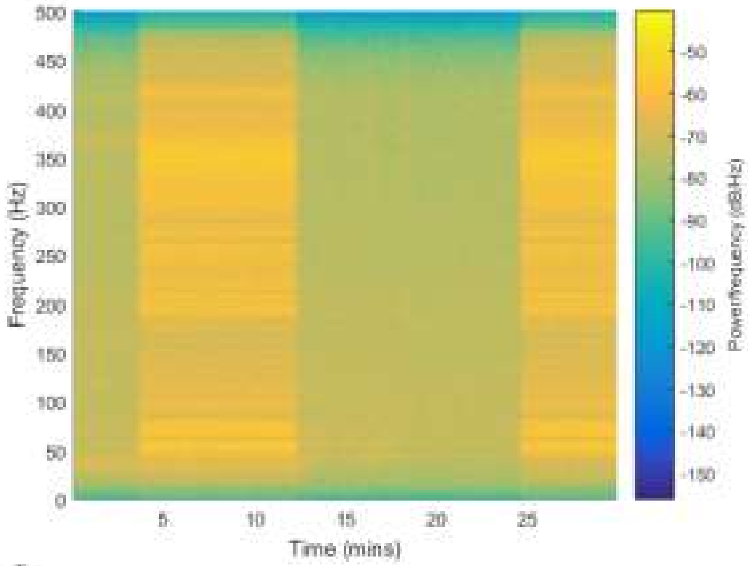

Fig. 3. Spectrogram of audio signal Train_Grid_A_A1.wav

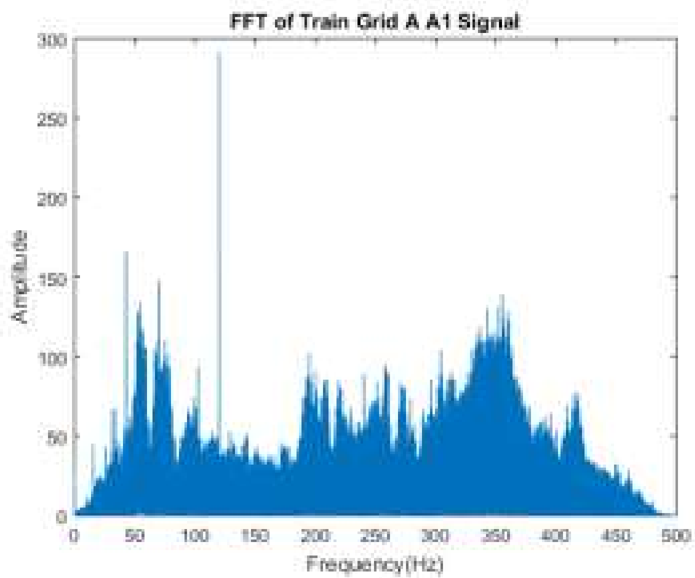

Fig. 4. FFT of audio signal Train_Grid_A_A1.wav

A. *Nominal Frequency Detection*

Figure 4 shows a FFT of an audio signal of nominal frequency 60Hz. But peak amplitude of the FFT is around a harmonic of 60Hz. This proves the requirement of considering both the base band of the nominal frequency and its harmonics for the detection of the nominal frequency. Basic approach was to compare the FFT mean amplitudes at 50 Hz and its harmonics with mean amplitude at 60 Hz and its harmonics. This method was observed to be quite sensitive to background noise. Therefore, more optimized method was designed to reduce the effect of noise.
But in this case, heights of the peaks relative to the neighboring

noise amplitudes were measured, instead of taking its amplitude from zero. To express maximum height about a frequency **f** quantitatively, a parameter **h** was defined as follows.

$$h = max(f - d1, f + d1) - \frac{mean(f - d1 - d2, f - d1) + mean(f + d1, f + d1 + d2)}{2} \quad (1)$$

The function max outputs the maximum and mean outputs the mean amplitudes of any frequency range given. Maximum amplitude between the frequencies *f - d1, f + d1* were found and mean amplitude of the frequency bands, *[(f - d1 - d2), (f - d1)],[(f + d1), (f + d1 + d2)]* were reduced from that value. **h** represents the strength of the frequency component **f**, relative to its surrounding frequency components.

Then h is calculated for 50/60 Hz and its harmonics. If the highest **h** value is generated for a multiple of 50Hz, the nominal frequency is detected as 50Hz; otherwise it is taken as 60Hz.

B. ENF Extraction

The proposed algorithm uses a Spectrogram based spectral combining approach for the ENF signal estimation. In the Spectrogram approach, first the given signal is segmented into smaller time frames and Short Time Fourier Transform (STFT) is applied to determine the spectrum corresponding to each time frame.

In the frequency domain ENF can be considered as a summation of impulses at the nominal frequency band and its harmonics. Here spectral components of base band and its harmonics are combined together to estimate the ENF signal. Additional frequency components which are interfering in the ENF signal as noise should be removed.

For a given time frame, the observed power spectrum component band will be,

$$P_{B,k}(f) = A_k h_k(f) + P_{n,k}(f) \quad (2)$$

Where $f \epsilon [k(f_0 - f_B), k(f_0 + f_B)]$ corresponding to a frequency band around kth harmonic as the nominal frequency. $P_{n,k}(f)$ represents noise component at the harmonic band. $h_k(f)$ denotes the spectral function of the ENF signal (Ideally an impulse-like function withpeaks at base band and each harmonic band.). In the algorithm for each time frame all these spectrum components are combined into the base band $[(f_0 - f_B), (f_0 + f_B)]$. Therefore the spectrum of the ENF signal at each time frame can be written as,

$$S(f) = \sum_{k=1}^{L} w_k P_{B,k}(kf) \quad (3)$$

The introduced w_k is a weight corresponding to each harmonic band, based on the signal to noise ratio (SNR) around the corresponding harmonic.

$$w_k = \frac{P_{Signal}}{P_{Noise}} \quad (4)$$

For the determination of spectral combining weights, take as an example and take $f_0 = 50, f_B = 1$. Therefore the band for calculating the ENF will be [49, 51]. Assuming that the ENF fluctuates in the band [49.6, 50.4], P_{Signal} will be the average power spectral density (PSD) within this band. P_{Noise} will be the average PSD in the bands [49, 49.6] and [50.4, 51]. This same method used to calculate the spectral combining weights of other harmonics as well.

ENF frequency $f_{ENF=f_0+\Delta f}$ can be obtained by finding the maximum of $S(f)$. It is possible to directly find the maximum in $S(f)$. But this leads much lesser accuracy since spectrogram calculates frequency for some finite number of discrete frequencies. In order to overcome this issue, maximum was estimated using Quadratic Interpolation [3]. But it was noticed that there are some outliers in the ENF signal, these outliers were replaced by the average of the samples before and after the given outlier.

III. Location Identification System

After successfully extracting the ENF signal from a multimedia signal(Power Recording or Audio Recording), classification of the signals should carried out. Classification was experimented using many efficient classifying algorithms such as Kth nearest neighbor classifier, nave bayes classifier, classification trees and neural network classifier. But as test results suggests, the most suited classifier for the above given task was the multiclass Support Vector Machine (SVM).

First step of the classification process is to train the classifier using both audio and power signals provided for nine grids. Each signal was sub-sampled to obtain 5 minutes long recordings in order to obtain more training data. Then ENF is extracted from each sub-sampled signal and the corresponding features were extracted from the ENF to train the classification system. The final accuracy obtained critically depends on the selected features of the signals. Therefore the task was experimented among different statistical features of the signal. In the final implementation following features were used for training of the classification system and for identification of a given signal of an unknown power grid [2].

- Nominal frequency
- Frequency range of ENF signal in a logarithmic scale
- Variance in a logarithmic scale
- Mean frequency shift from nominal frequency (Difference between nominal frequency and mean frequency).
- Variance of the approximate signal at the 7 th level (in a logarithmic signal) with the wavelet transform for 7 levels
- Variance of the detailed signal (in a logarithmic scale) with first five levels of the wavelet transform for seven levels

In the given design task there are two types of signals that are given. They are ENF extracted from either power recordings or audio recordings. Generally ENF extracted from Power signals are cleaner with a higher signal to noise ratio (SNR) whereas Audio recordings which are normally mixed with noise have a lesser SNR value [2]. The final implemented

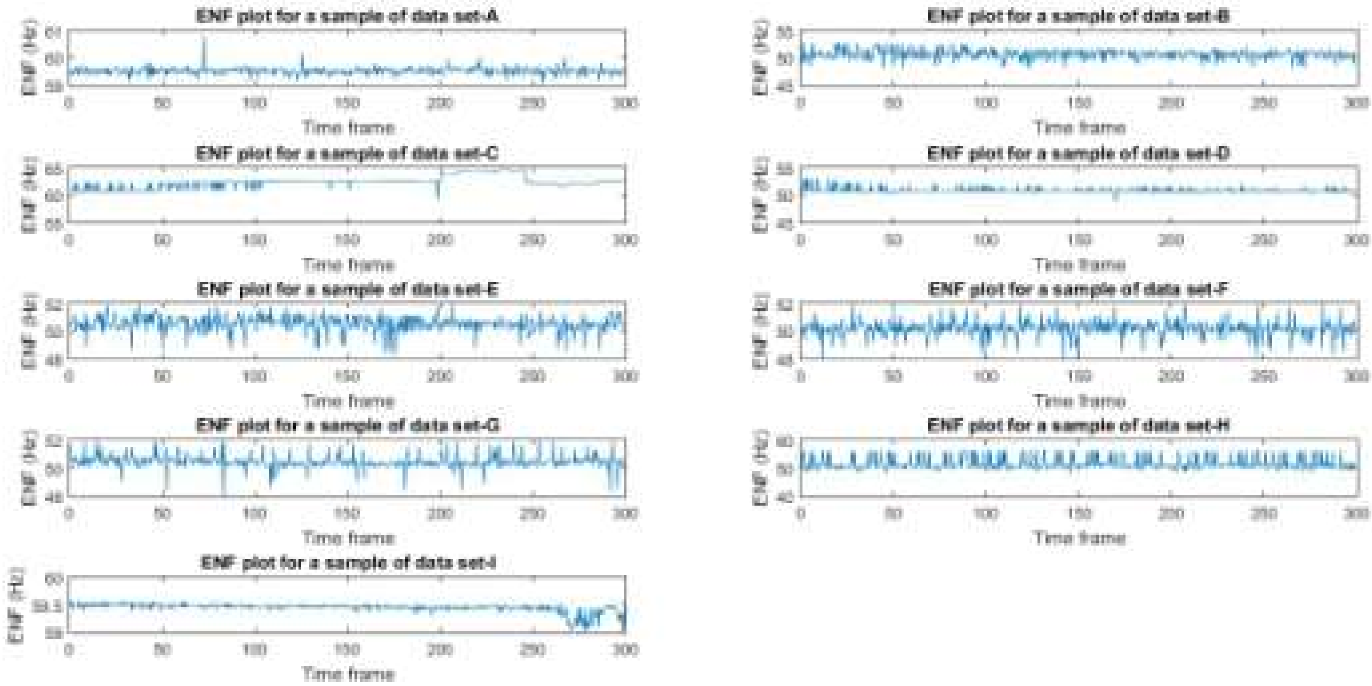

Fig. 5. ENF extracted from Audio signals

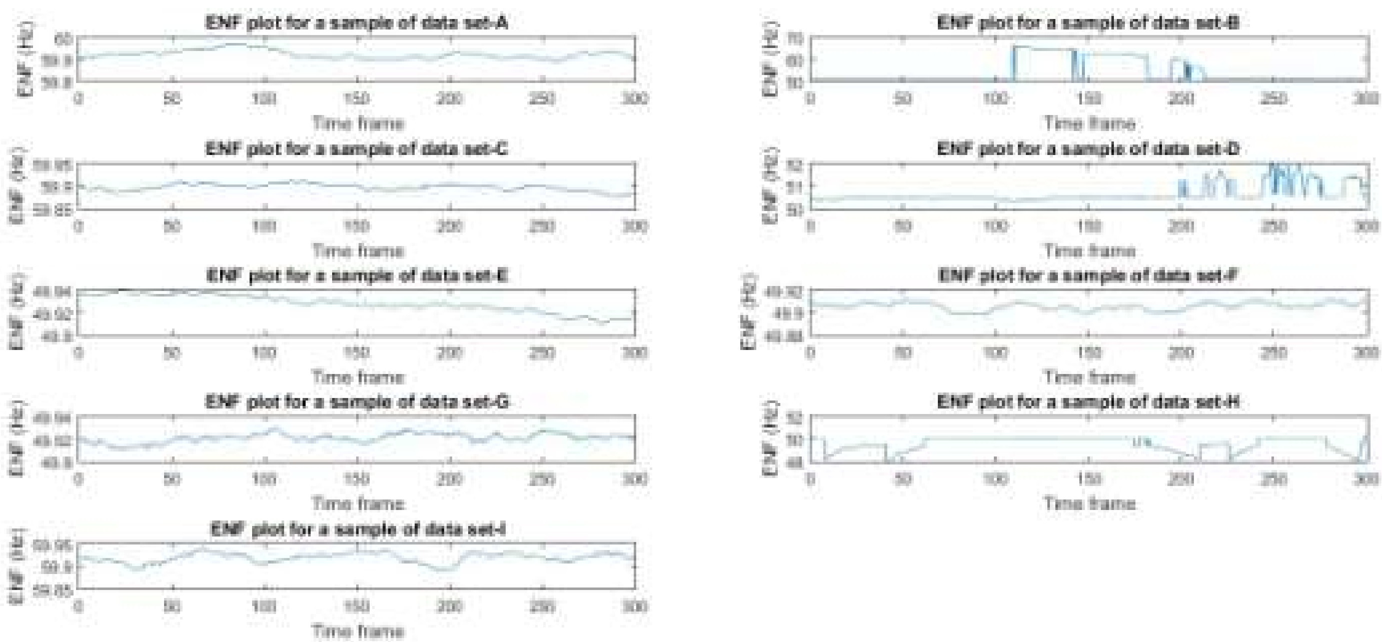

Fig. 6. ENF extracted from Power signals

classification system is trained with ENF extracted from both Audio and Power signal recordings.
Signals, inserted for training, are divided into two categories according to the nominal frequency of the ENF signal. Then two separate classifiers were trained using rest of the features, one from the features extracted from 50 Hz nominal frequency signals and the other one from the features extracted from 60 Hz nominal frequency signals. So in the final classification system, there are two trained models separately for 50Hz and 60Hz nominal frequencies. As mentioned above, multiclass Support Vector Machine (SVM) is used as the classification algorithm for the final classification system.
When a signal is inserted for grid classification, above listed features were extracted from its ENF and fed into the corresponding trained classifier model depending on the nominal frequency of the ENF. Then probability scores for each grid are taken from the system. A threshold was used to determine whether that particular signal belongs to a given grid or ′None of the above′.
In the final classification system a multiclass Support Vector Machine (SVM) with a quadratic polynomial is used as the kernal function.

IV. Results & Discussion

A. Results for Practice Dataset

Practicing dataset classification accuracies for various classification systems are listed in the table 1.

Classification Algorithm	Accuracy for practicing dataset
K th Nearest Neighbor	52%
Neural Network	38%
Bagged Tree	66%
Gaussian SVM	58%
Quadratic SVM	72%

TABLE I
Practice Dataset Accuracy

Final grid classification result for practicing data set:
AHCFF,FGIID,AFHDC,IIAAE,BBAD,CHFGB,DFCHG, EACHI,BHECF,BAGGI

Final grid classification result for testing data set:
IGDID,FFDAF,AAGBG,FFCEH,BHHIB,BFGAI,DIFHN, IBCBH,EIIBE,FGAAG,CIIIE,HAEBC,NCFDG,CECBI,EICDA, BBBEA,GINIG,AABIH,CADBA,NBFBB

As the results suggest highest accuracy for practice dataset was obtained for quadratic SVM which is 72%.
A neural network was also used as an approach, but the accuracy was comparatively lesser since it requires much more training data for the classification.
As another approach, if the confidence value is lower than a certain threshold, the system was advanced into a second stage. Here the features of the ENF are subjected to a binary classifier trained on the two classes which received the highest confidence values in the first stage [2]. But this approach did not improve the accuracy as expected.

V. Circuit design

Task of the circuit was to record a power signal for at least for 30 minutes from a standard power outlet. An arduino development board was used to record the step downed power signal. Collected data was transferred to a computer using a serial cable and converted to a .wav file using MATLAB.

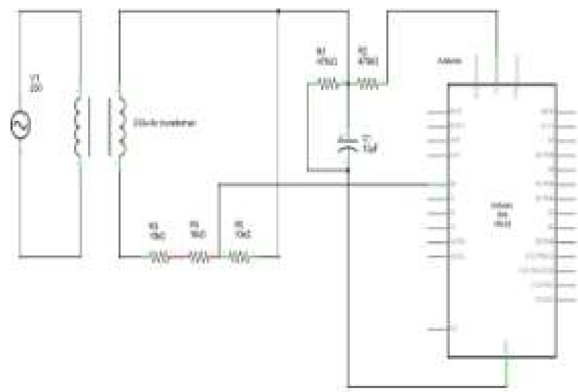

Fig. 7. Circuit Schematics

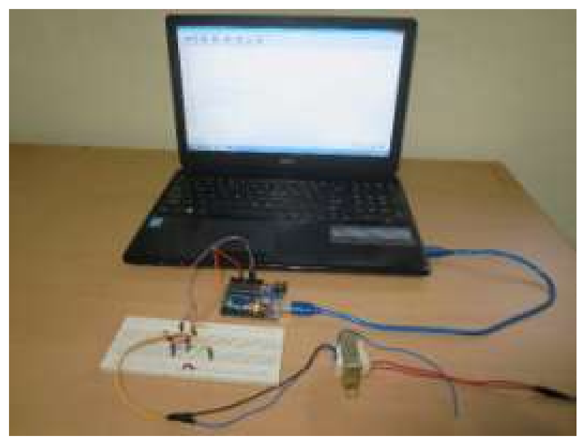

Fig. 8. Photo of the Circuit

A. Circuit Overview

- 230 V AC voltage was step down to 6 V, using a transformer.
- 6 V AC voltage again stepped down to 2 V, using three 10k resistors (voltage divider).
- USB connection was used for both the serial connection and the power supply for the Arduino.
- 5 V reference voltage was taken from Arduino and divided into two parts of 2.5 V each, using two 470k resistors (voltage divider).
- Capacitor between 1uF - 10uF was used between ground and 2.5 V point to stabilize the voltage.
- 2.5 V DC voltage was used as an offset to the 2 V AC signal.
- Output of 2 V peak to peak AC signal with 2.5 V DC offset was given to an Arduino analog pin.

B. Arduino Code

- Timer interrupt of the Arduino was used in CTC mc (Clear Timer on Compare match) with a pre-scaler 256.
- Compare register was given a value of 62 to get the 1C Hz frequency needed for updating the register.

$$Clock\ Cycles\ needed\ for\ timer\ interrupt$$
$$= (16MHz/256) * 2000 \approx 62$$

62 was chosen as the register value considering delays the real time functioning.

- Analog to Digital Converter (ADC) of Arduino was us approximately at 1 KHz frequency.
- Baud rate of the communication link between compu and Arduino board was set to 230400. In lower Ba rates than above value did not work properly due interference of interrupts with communication rate (ju sample values were received in lower Baud rates).

Data coming from Arduino were acquired (saved to a .txt file) using a software called CoolTerm. These .txt files were read using MATLAB and saved as .wav files after normalization.

VI. Analysis of Recorded Signals

Twenty power signals were recorded, each with 30 minutes duration, add up to 10 hours of recorded signals. After the analysis of the recorded signals using the implemented algorithms it was observed that the nominal frequency of the grid is 50Hz. Extracted ENF of two recordings from different times of a day is shown in the figure 9 and 10.

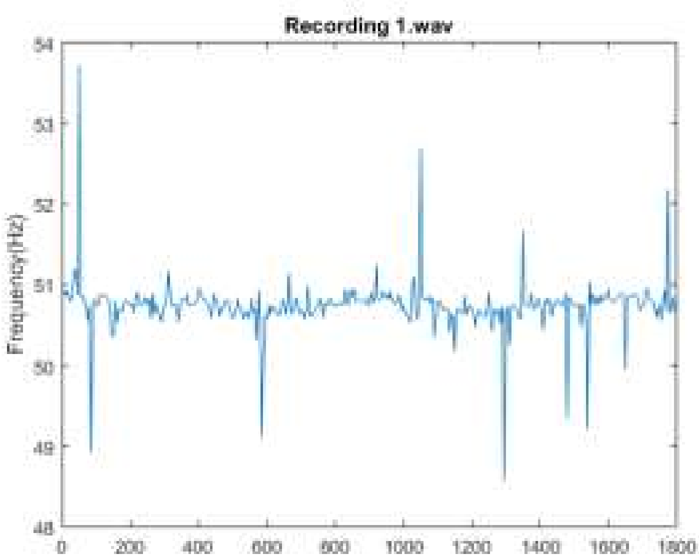

Fig. 9. Extracted ENF of Recording_1.wav

The features obtained for 5 of these Power Recordings are shown in the figure 11. According to the extracted ENF signals, variances of the ENF signals are quite low compared to some of the given grids.

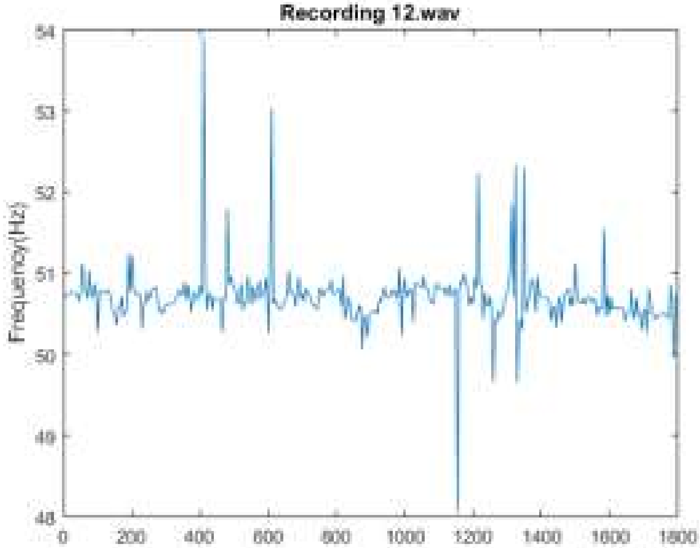

Fig. 10. Extracted ENF of Recording_12.wav

Recording	Mean difference	Log(Range)	Log(Variance)	Log(Variance) of Approximate Signal of Wavelet	Variances of first 5 levels of Wavelet transform (Detailed Signal)				
Recording 1	0.7677	-1.6064	-6.8544	-2.3733	-14.7401	-13.6732	-11.0289	-8.0891	-5.7716
Recording 2	0.7023	-2.1199	-7.7768	-3.3342	-12.9992	-12.4885	-10.8118	-8.6312	-6.4544
Recording 3	0.7671	0.5344	-3.3847	-0.8021	-6.1422	-5.0852	-3.6255	-1.9203	-1.8897
Recording 4	0.7531	-1.7115	-7.3273	-2.7689	-13.5251	-12.3129	-10.6432	-7.8399	-6.3367
Recording 5	0.7301	0.2788	-5.3782	-2.3167	-7.6168	-6.3080	-5.0113	-3.9605	-38686

Fig. 11. Extracted features for power recordings.

Mean of the ENF signals varies around 50.7 Hz. There are sudden spikes in the frequency of ENF signals. This can be clearly observed in figure 9 and 10.

VII. Conclusion

The basic task of this project was to design a system which can identify the captured location of a given multimedia sample based on the ENF signal embedded in it. Identification of the Nominal frequency was the initial step in detecting of ENF. Once Nominal frequency was found, ENF was extracted using a spectrogram based spectral combining approach. Then a classification system was trained with some statistical features extracted from ENF signals such as nominal frequency, frequency range of ENF signal and variance in a logarithmic scale. Finally, a trained model was obtained which was used for the classification of ENF signals of unknown power grids. After experimenting with several classification algorithms including SVM, KNN (K^{th} Nearest Neighbour), Neural Network etc. highest overall accuracy obtained for the practice dataset

was 72%. For the final implementation, multi-class Support Vector Machine with quadratic kernel function was used.
Second task was to build a hardware device to record ENF from a standard power outlet. The device which was built using the Arduino platform, was successful in obtaining power signals from the domestic supply.

ACKNOWLEDGMENT

We would like to thank our supervisor Dr. Chandika Wavegedara, 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 (September 2013):885-888

[2] Adi Hajj-Ahmad, Ravi Garg, Min Wu.*ENF-Based Region-of-Recording Identification for Media Signals.* vol 10,no. 6 (June 2015):1125- 1136

[3] Adi Hajj-Ahmad, Ravi Garg, Min Wu. *Instantaneous Frequency Estimation and Localization for ENF Signals.*

[4] Math H.J Bollen, Irene Y.H Gu. *Signal Processing of Power Quality Disturbances.* IEEE Press Series on Power Engineering.