Exploring Power Signatures for Location Forensics of Media Recordings

The Zenith
SP Cup Team ID: 24514
College of Engineering,Trivandrum,India

***Abstract*—This report presents the results to the challenge of "Exploring Power Signatures for Location Forensics of Media Recordings" as apart of Signal Processing Cup 2016 by IEEE Signal Processing Society. Here we examine different frequency estimation and classification techniques to provide accurate ENF estimates and classify these signals into the corresponding grid of recording. In this report we propose methods of efficient extraction of ENF signal using quadratic interpolation and frequency tracking.The SVM and GMM classifiers used provided a classification accuracy of 80% on practice data set.The result for testing dataset is also given.**

I. Introduction

Electric Network Frequency (ENF) is the supply frequency in power distribution grids. It has a nominal value of 60Hz in North America and 50Hz in Europe, most of Africa, Asia and Australia. The ENF fluctuates around its nominal value because of load variations and defects in control mechanism of the grid.The variation of supply frequency with respect to time is called an ENF signal.

ENF signals are extracted directly from power recordings obtained using a signal recorder which is fed with a stepped down and voltage divided signal from power outlet. The recorded signal is split into time frames of equal duration and different frequency estimation techniques are used to estimate the prominent frequency around the nominal ENF value. Combining these instantaneous estimates gives an estimate of the ENF signal. The ENF fluctuation are also present in audio recordings made using devices in areas under the proximity of an electrical activity [4]. This is due to electromagnetic influences at the place of recording. It has been shown that the ENF variations extracted from the same power grid are similar which can be used to uniquely to identify the grid from others[1].

ENF signal based analysis has paved way for multiple forensic applications [9]. It has been shown that the ENF can be used to detect tampering or modification in a multimedia signal. Estimation and validation of time of recording and also its location across the grids are popular application of ENF signal. However, such application requires a vast database of synchronous power signal with the knowledge of grid of origin which is used as reference.

In the report we explain in detail the ENF extraction analysis and classification strategy followed to identify the grid of recording. Different methods for extracting the ENF signal from power or audio are presented in [5,13,15].Given a database of such recordings corresponding to certain power grids, different approaches for classifying them are described in[17][18].We studied various ENF extraction techniques and some of the techniques we used to solve the challenge are briefly described in the following section of this report.We also propose a modified frequency tracking approach by defining a new cost function.

The remaining part of the report is organised as follows. Section II briefly describes the theory behind the different methods used for extracting ENF signal from the recordings along with a proposed Frequency Tracking Algorithm. Section III presents the classification system used and the proposed features to describe the extracted ENF signals. Section IV discusses the experiments done on the practice and test dataset. Section V examines the circuit design used for recording a power signal and analysis of ENF extracted from the recording of the current location with that the 9 grids provided and Section VI concludes the report.

II. Extraction of ENF Signals

The ENF extraction techniques that we studied and implemented as part of the SPCUP 2016 are :
A.Short Time Fourier Transform
B.Quadratic Interpolation
C.Improved DFT Approach
D.Frequency Tracking Method
E.Proposed Modified Frequency Tracking and
F.Estimation of Signal Parameters via Rotational Invariance Technique(ESPRIT).

We now present a brief description of the above techniques.

A. Short Time Fourier Transform

A basic approach for extracting ENF would be to make use of the Short Time Fourier Transform. This approach utilizes the distribution of power among various frequencies in small time durations, here after mentioned as a "time frame". The signal is assumed to be stationary in each individual time frame. For each time frame, the Discrete Fourier transform (DFT) of the signal is computed and the frequency bin with maximum power component is chosen as the estimate of the ENF for that single time frame.

In order to get a good estimate of frequency (i.e better frequency resolution), one would have to take as many samples of the signal as possible which in turn would limit the time resolution, which is determined by the duration of time for which we assume the signal to be stationary. Thus a frame size has to be fixed considering the trade off between time resolution and the frequency resolution, a difficult task in practical scenarios. To overcome this limitation, adjacent time frames over which the ENF is estimated is made to overlap with each other, thereby reducing the time for which the signals is assumed to be stationary. Overlapping may sound perfect at first but it unnecessarily increases computation overhead. Even with overlapping, the short time Fourier transform approach still lags in terms of performance in low SNR signal estimation.

B. Quadratic interpolation at maximum power

This is similar to the STFT based technique but uses a quadratic interpolation to increase the frequency accuracy. In this, the signal is split to blocks of length N with 50% overlap and the power spectrum of the blocks is computed. For each block, the frequency bin b_{max} corresponding to the maximum power is noted and a quadratic model to the frequencies at $b_{max}-1$, b_{max} and $b_{max}+1$ is fitted. As described in [5], we compute the value of parabola peak p as:

$$p = 0.5 \times \frac{(y(b_{max}-1) - y(b_{max}+1))}{(y(b_{max}-1) - 2y(b_{max}) + y(b_{max}+1))} \quad (1)$$

where,

$$y(b) = 20log_{10}(PSD(b))$$

and PSD(b) represents Power Spectral Density at frequency b. The ENF of the block is then given by:

$$F_{ENF} = \left(\frac{b_{max}+p}{N}\right) F_s \quad (2)$$

where F_s is the sampling frequency of the signal.

A detailed description of this method can be found in [5].

C. Improved DFT Approach

In this method [15], a better approximation of the Discrete Time Fourier Tranform (DTFT) of each N length block is achieved by increasing the number of FFT points from N to a very high value ($\sim 1000N$). As explained in the STFT based method, the ENF is assumed to be the frequency at which the amplitude of the FFT is maximum. Due to the discrete approximation of the DTFT by the FFT, the ENF estimated by this method can be different from the actual value. This is the rationale behind increasing the number of FFT points which will help to better approximate the DTFT of the block.

It is impractical to compute such high point FFT for each block. Instead, initially the N point FFT of the block is computed to determine the frequency F_{ENF} corresponding to the maximum amplitude. With this value of F_{ENF}, Fourier transform for a small part of the spectrum centered around F_{ENF} is only computed. The idea is to estimate power in the region between the bins adjacent to the F_{ENF} bin by using the appropriate set of basis functions which are generated from the kernel of a high resolution DFT. The given signal for which the spectrum is to be recomputed is projected onto the higher resolution bases and square of this projection value is taken as the power distributed in that frequency bin. Once the spectrum component is estimated, a binary search is performed for estimating the local maxima in the spectral distribution.

The search is done until the middle bin is found to have a higher power density than the adjacent bins. This limits the search to a local maxima (we only expect one maxima).

D. Frequency Tracking Method

Common non parametric approaches like quadratic interpolation at maximum power and the short time Fourier transform described above take into account only those frequencies that contribute maximum to the spectral power and therefore tend to be inaccurate at low SNR levels. In this scenario, the signal of interest should be characterized by using properties other than the maximum power parameter. Following this line of thought, it has been proposed in [13] that the property of low variance of ENF signal over time be used to distinguish it from noise. The method aims to minimize the squared difference of frequency estimates over adjacent time frames, thereby obtaining a local minimum for the overall ENF signal variance. The algorithm[13] minimizes the following cost function to obtain the minimum variance ENF estimate

$$J(j, f_j) = \sum_{r=j+1}^{R} (f_r - f_{r-1})^2, \qquad f_j \in \lambda_j \quad (3)$$

where f_j and λ_j are the frequency estimate and the set of frequency bins respectively in the j^{th} time frame.

The approach of iteratively fixing the optimum value of λ_j is computationally inefficient. Therefore, we choose a predetermined, static number of bins for each time frame and propose a new cost function that also accounts for the power in the frequency bin in addition to the frequency jump from the previous bin.

E. Proposed Modified Frequency tracking

An optimum solution to the estimation problem can be obtained by maximising the power contained in the signal and simultaneously minimising the overall variance of the estimated signal.

To obtain the optimum solution, we propose the following cost function

$$J(f_j) = \|f_j - \bar{f}\|^2 + k(1 - P(f_j)), \qquad f_j \in \lambda_j \quad (4)$$

$$\bar{f} = mean(f_1, f_2, \cdots, f_{j-1}) \tag{5}$$

where $P(f_j)$ is the power in the j^{th} frequency bin of the R^{th} frame, given by

$$P(f_j) = \frac{I_R(f_j)}{\max\limits_{r \in \lambda_R} I_R(f_r)} \tag{6}$$

and k is the weight penalising the power content in the signal (or more precisely, the absence of it).

In order to find a global minimum for this cost function, 'λ' frequencies at which the power spectrum peaks is found out for each time frame and the cost function is computed for each frequency jump to adjacent frame. Forward and backward costs are computed and for the backward jump from the first frame, cost is assumed to be infinite. The path which minimises the cost function in the forward path is chosen if the cost is below a pre-adjusted threshold, if not, the backward path is chosen until all frequency peaks for the previous frame has been tried out. In such a situation, the threshold is reduced and the whole process repeated. If not, the algorithm picks off from the new estimate and repeats the entire process.

The algorithm employed, popularly known as the Fano Algorithm[19], flow chart of which is given in Figure 1,is used for decoding convolution codes and is cherished for its low memory requirement and lenient computational overhead.

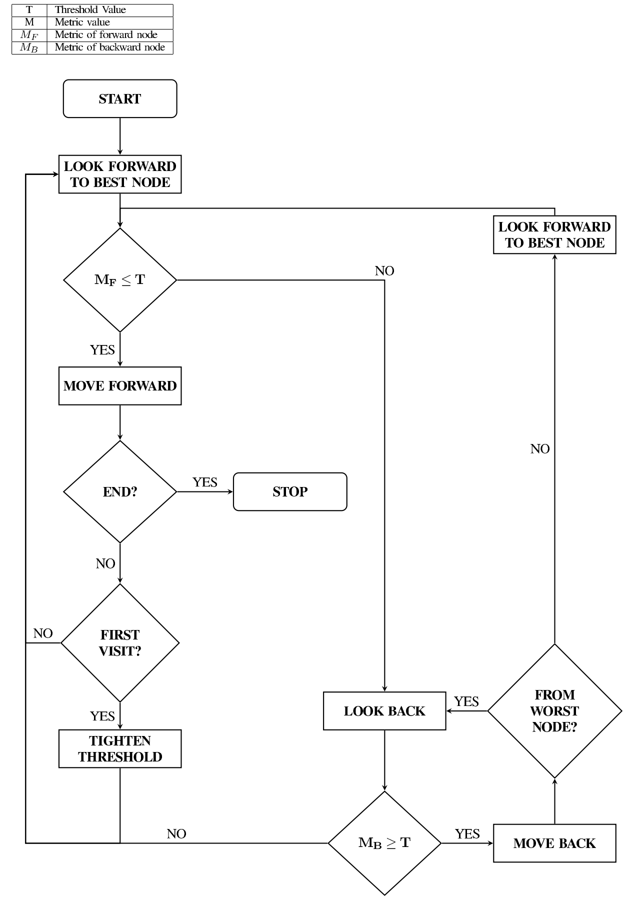

Fig. 1: Flowchart of Fano algorithm.

F. Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT)

The parametric subspace based approach for frequency estimation is more precise because it makes prior assumption of the signal to be estimated. ESPRIT[12] is the most widely used parametric frequency estimation method based on the subspace analysis of a signal and noise model. In general, these methods can be used to estimate the frequency of a signal composed of P complex exponential embedded in white noise W_n.

$$x_n = \sum_{i=1}^{P} A_i e^{jn\omega_i} + W_n \tag{7}$$

where A_i refers to amplitudes of the sinusoids with frequencies w_i. As ENF signals consist of only one real sinusoid, the value of P for ENF signals is 2. ESPRIT makes use of the rotational property between staggered subspaces that is invoked to produce the frequency estimates. In our case, this property relies on observations of the signal over two intervals of the same length staggered in time. The sample correlation matrix X from data is computed which is Singular Value Decomposed as

$$X = LSU^H \tag{8}$$

Where L is an $N \times N$ matrix of the left singular vectors, S is an $N \times M$ matrix with its main diagonal entries containing the singular values, and U is an $M \times M$ matrix of the right singular vectors. The singular values correspond to the square roots of the eigenvalues of the sample correlation matrix X scaled by length of data matrix N, and the columns of U are the eigenvectors of X. These vectors form an orthonormal basis for the underlying M-dimensional vector space. More specifically, U can be written as U =$[U_S|U_N]$, where U_S is the $M \times P$ matrix of right singular vectors corresponding to the singular values with the P largest magnitudes and U_N is the $M \times (M - P)$ matrix containing the remaining right singular vector. The signal subspace can be partitioned into two smaller (M - 1) dimensional subspaces as:

$$\mathbf{U_S} = \begin{bmatrix} \mathbf{U_1} \\ * \end{bmatrix} = \begin{bmatrix} * \\ \mathbf{U_2} \end{bmatrix} \tag{9}$$

Where $\mathbf{U_1}$and $\mathbf{U_2}$ correspond to the unstaggered and staggered subspaces, respectively. The Relation between $\mathbf{U_1}$and $\mathbf{U_2}$ can written as: $\mathbf{U_2}$=$\mathbf{U_1}Q$ Where Q is a $P \times P$ matrix. Q

can be computed using least squares method. Eigen analysis can then be carried out on Q. The frequency estimates can be extracted from the arguments of the eigenvalue values of Q by ϕ_k for $1 \leq k \leq P$, the frequency estimates are given by

$$\hat{f}_k = \frac{\angle \phi_k}{2\pi} \qquad \text{with} \qquad 1 \leq k \leq P \tag{10}$$

A detailed description of the algorithm can be found in [13].

III. Multi Class Classification

Once the of ENF signals are extracted from the given collection of power and audio recordings of different grids, our next task is to classify them into appropriate grids. To this end, first we extract some distinguishing and non-redundant features from the ENF data followed by training a multi-class classifier with these features.

A. Feature Extraction and Analysis

To maximize the number of training samples, feature extraction was done on sub blocks of the estimated ENF signal. In addition to the 16 length feature vector used in [11], we propose a set of 6 additional features for better classification. Our 22 feature set can be used to improve classification.

Mean, variance and higher order moments are important statistical parameters and therefore taken as features.

The Wavelet transform, which represents a signal as a projection onto basis vectors with varying time-frequency resolutions, can be used as a very good descriptor of the signal. Wavelet decomposition of a signal generates detail and approximation coefficients. The approximation coefficients represent the relatively invariant component of the signal across time frames. The detail coefficients represent the quick variation of the ENF signal in subsequent time frames. The variance of these coefficients at each level accurately describes the signal.

The signal is modelled using a M level auto regressive function to represent the correlation among adjacent samples. The innovation signal $v[n]$ is a measure of the convergence of the model.We use an AR(5) model given by

$$s[n] = a_1 s[n-1] + a_2 s[n-2] \cdots + a_5 s[n-5] + v[n] \tag{11}$$

TABLE I: Feature Components

Index	Features
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 nine levels of detail signal computed through L-Level wavelet analysis from coarser to finer(L=9).
14-17	$AR(5)$ model parameters α_2,α_3,α_4 and α_5 .
18	log(variance) of the innovation signal after $AR(5)$ modelling.
19	Variance of power spectral density
20-21	Higher order moment 3 and 4
22	Variance of cross-correlation between adjacent time frames

The coefficients are normalised by dividing each with the first coefficient a_1. The four normalised AR coefficients resulting from modelling a_2, a_3, a_4 and a_5 and the variance of the model's innovation signal $v[n]$ are considered as features from the AR model.These features have the potential to distinguish ENF signals in terms of how well they can fit such an autoregressive model and in what manner.

Proposed New Features: We find the power spectral density of the ENF signal and find the variance of this frequency domain signal and used as an additional feature . This measure captures the variation of the ENF signal in the frequency domain.

A slowly varying ENF signal would be highly correlated with itself and would show lower variation between adjacent samples. Therefore variation of cross-correlation between adjacent time frames is also used as a feature.Thus we propose 6 additional features compared to [11] which are variance of power spectral density, variance of cross-correlation and third and fourth higher order moments and the two additional coefficients for the AR(5) model . The feature components that we use for location classification are summarized in TABLE I.

We take log of the range and variance feature values to emphasis on their orders of magnitude and potentially enhance the separability between the final feature values. Normalization is done to increase the separation and distinguish each grid from each other.The computed feature values are normalized to the range of [-100, 100] by a linear scaling, whereby the k^{th} feature value in a training example is normalized according to the other feature values in position k in all training examples. The normalization parameters are stored and later applied to the testing examples to normalize them. We follow the normalization procedures as in [2,11].

B. SVM Classifier

SVM is basically a two-class classifier based on the idea of large margin and mapping data into a higher dimensional space. The principle of a SVM is to construct a hyperplane or a set of hyperplanes in a high or infinite dimensional space, which can be used for classification so that a good separation is achieved by the hyperplane that has the largest distance to the nearest training-data point of any class. In general, the larger the margin, the better the classifier.

When the feature space is not linearly separable, SVM maps the data into high dimensional feature space with non-linear mapping, and finds the optimal classification hyperplane in the high dimensional feature space.

After extracting features from a raw ENF signal . We trained the system using the one vs one approach which uses a group of binary classifiers. The SVM implementation produces ${}^{m}C_2$ binary classifiers for M classes. Each binary classifiers are trained on one of the possible pairs from the M grids, which learns to distinguish between the respective two grids. During the testing of a trained classifier system, we give each test vector through each binary classifiers. Based on the results from the binary classifiers, LIBSVM assigns scores to each grid. The final answer is the grid with the largest value of score. LIBSVM also gives probability estimates which is a measure of its proximity to the cluster corresponding to that grid[11].

There is a mismatch in the number of recordings from different grids. If a system is trained on a dataset where the majority of the training examples belong to one class, it tends to be more biased to the class with highest number of training samples. To overcome this issue, we use a variant of SVM called the weighted SVM, which is supported by LIBSVM. The weighted SVM addresses the issue of imbalanced data through assigning different cost values for examples from different classes. The larger class has a smaller cost value than the smaller class, which means that the penalty for making a mistake on an example from the smaller class would be larger. Here, with M classes, the cost for class j that has N_J training examples would be $w_j \cdot C$ where

$$w_j = \frac{N_{min}}{N_j} \qquad for\ j = 1, 2, 3 \cdots, M\ and\ N_{min} = \min N_j \tag{12}$$

In our implementations, we use the linear Function kernel for our SVMs. Using the LIBSVM library for the linear kernel, the cost parameter C can be controlled which relates to how far the influence of a single training example reaches. For each SVM classifier we train, we select the value for C through cross-validation. The parameter C is a compromise between the minimum distance from the separating hyperplane and the number of outliers possible in the model

IV. Experimental Results

We now present the details of the experiments performed on the practise dataset. Based on these experimentations we select the best ENF extraction technique and also develop an optimum classification strategy and these optimum ENF extraction- classification scheme is applied to the test data set and report our final result.

A. Training data

The training dataset provided, consisted of a total of 111 signals of varying durations of 30 minutes to 90 minutes, collected from 9 different grids, of which 93 were power recordings and 18 were audio recordings. All these recordings were either sampled at or down sampled to (after low pass filtering using proper anti-aliasing filters) 1000Hz. The practice dataset and the testing dataset provided consisted of 50 and 100 signals respectively, each of 10 minutes duration, sampled at 1000Hz.These were not labelled as audio or power recordings and could have been recorded from one of the above mentioned 9 power grids or may not be from any of these grids. The spectrogram of a typical power signal is shown Figure 2.

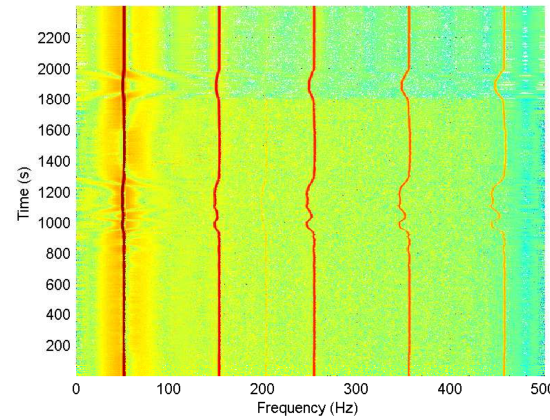

Fig. 2: Spectogram of 60 Hz Power signal

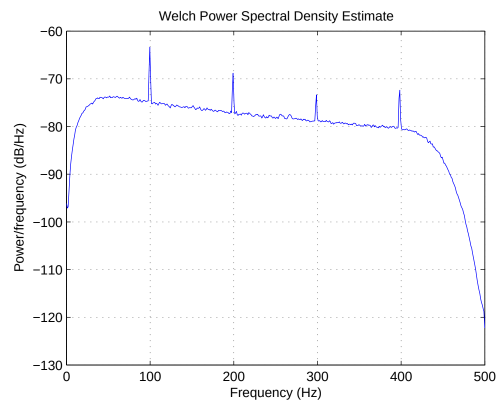

Fig. 3: Power spectrum of 100 Hz Power signal

B. Pre-processing

As the nominal value of ENF is either 60/50Hz, we down-sampled the signals to 500Hz, reducing computational complexity, without any loss of information. We chose 500Hz to make possible ENF extraction at one of its higher harmonics if needed. To determine the nominal value of enf, we computed the power at the fundamental frequncies (50 and 60 Hz) and their next three harmonics and compared the sum of powers .

Since the ENF signal is only present in a narrow frequency band centered at its nominal frequency, we applied a band pass filter centred at the nominal ENF frequency (or one of its harmonics) with a bandwidth of 3Hz.

C. ENF extraction

When we analysed the power spectral density plots of the power and audio recordings, we found that the ENF signal had maximum strength at 50/60 Hz for most of the signals, but for some recordings it was observed that the strength of the ENF was higher in one of its harmonics. A recording for which this is true is shown in Figure 3. This prompted us to extract ENF at the harmonic where its power is maximum. Therefore, for the signal in Figure 3, ENF would be extracted from the second harmonic (100Hz).

We implemented the six ENF extraction methods described in section II. The specifics and results for each are presented below.The results are presented in the increasing order of accuracy during classification :

1) STFT: In this extraction technique, we divided the bandpass filtered signal into overlapping blocks of size 1024 with overlap factor 50%. After this we computed the 1000 point FFT of each block. ENF frequency in each block is the frequency corresponding to the global maxima of its FFT.

Figure 4 shows a plot of ENF extracted using this method. We compared the ENF extracted by this method to typical ENF plots and found it to be severly handicapped in estimating the ENF. The main problem was frequency resolution. The ENF extracted by this method jumped between different values after small time intervals.

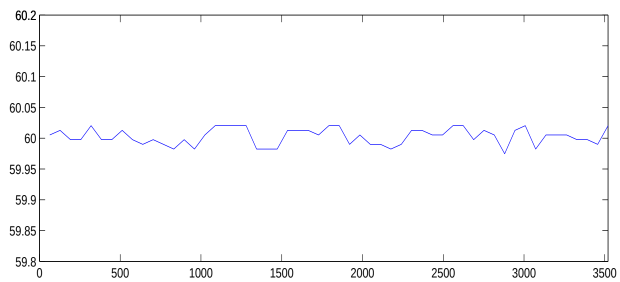

Fig. 4: ENF extracted from power signal using STFT.

2) Improved DFT aproach: For this method, we compute the 1024 point DFT for the signals at fundamental frequency and then improve the resolution further as described in section II C. The result of improved DFT method of ENF extraction is given in Figure 5. We found the improved DFT approach to give high resolution ENF estimates but it suffered from intermittent spikes.

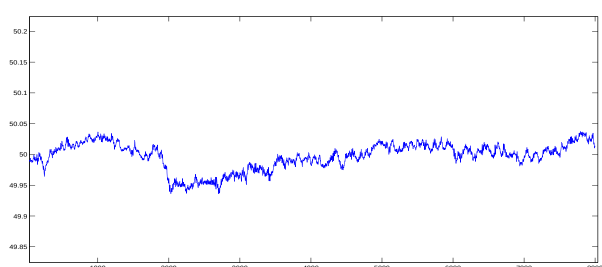

Fig. 5: ENF extracted using Improved DFT method.

3) ESPIRIT: Here we split the signal into blocks of 1 second duration with 50% overlap. We applied the ESPIRIT algorithm (with paramters in section II) to each block which gave us the ENF frequency in it.

The ENF extracted is shown in Figure 6. This method had high computation time though it gave high resolution estimates.

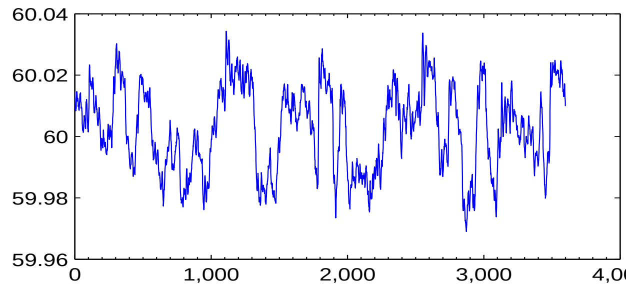

Fig. 6: ENF extracted from power signal using ESPRIT.

4) Quadratic Interpolation at maximum power: In this, the signal was split into blocks of 512 length each with 50% overlap. The frequency corresponding to maximum power and its adjacent frequencies were determined .

The result of ENF extraction by this method is shown in Figure 7. This method seemed to give relatively good estimates of ENF but we observed an offset of 1Hz or so from the nominal ENF frequency in some cases.

5) Frequency Tracking: The frequency tracking algorithm was applied on time frames consisting of 1024 samples, from which 5 frequency bins at which the power peaked and the corresponding power content were stored. After storing these values for the entire time frame, ENF signal was chosen as that spectrum component which minimised the cost function (3). A plot of ENF extracted by this method is shown in Figure 8.

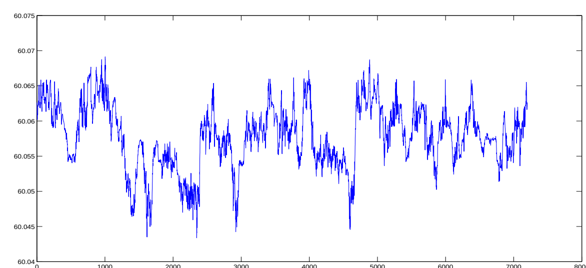

Fig. 8: ENF extracted from power signal using Proposed Modified Frequency Tracking.

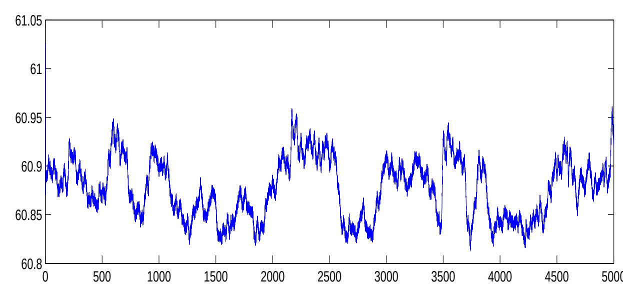

Fig. 7: ENF extracted from power signal using Quadratic Interpolation.

6) Proposed Modified Frequency Tracking: Here we show the results of the proposed computationally efficient modified frequency tracking approach.The new cost function function reduced the computational time for extracting the same ENF signal by 30%.

Occasional spikes were observed in some cases. We replaced these by the average of the previous two values. This method proved to be more robust in estimating ENF signal in the presence of noise as evident from the correlation coefficient comparison as shown in Table II.

TABLE II: Correlation Coefficients of Algorithms

SNR(dB)	Quadratic Intepolation	Proposed Frequency Tracking	ESPIRIT
10	0.5332	0.7242	0.9528
15	0.7284	0.8818	0.9668
20	0.8879	0.9575	0.9658
25	0.9593	0.9860	0.9880

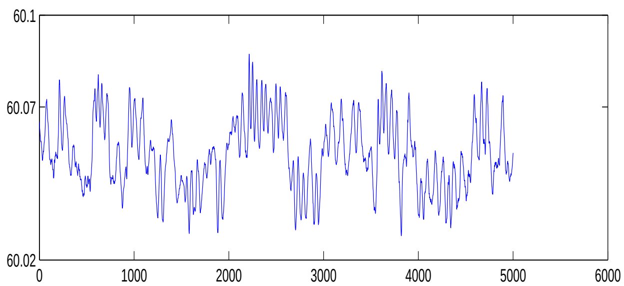

Fig. 9: ENF extracted from power signal using Frequency tracking.

D. Classification

We generated five different SVM models using the features extracted from the ENF estimated by different techniques as above. The various features we tried are shown in Table I. For judging how good a feature is, we plotted a series of histograms depicting the variation of features within a class and between two classes as shown in Figures 10 and 11.

After trying various combination of features for each model, we found that there is an optimum set of features which gives maximum classification accuracy on the practice dataset. This set of features varied across each extraction method. The result of accuracy obtained on the practice data set for each extraction method with the optimum set of features found for it is tabulated in Table III. The proposed modified frequency tracking approach and quadratic interpolation methods gave the highest accuracy with the features 1-22 and 1-16 (Table I) respectively.

Since both these methods used very different set of features and the signals misclassified them appeared to be different, we had the idea of combining these two models in some manner.

The approach we took to combine the two models was to add the weighted probability matrix output by the SVM for each model, and then find the class which had the maximum probability. For the probability matrix generated by each model, we took the inverse of the variance of the probabilities of a signal belonging to each grid predicted by the model as its weight.

Using this method, we got an improvement of 6% in accuracy making our final prediction accuracy as 72%.

TABLE III: Accuracy of Classification Models

Sl.No.	Extraction Method	Classification Accuracy Obtained
1.	STFT based method	34%
2.	Improved DFT	32%
3.	Quadratic interpolation	66%
4.	Frequency tracking	68%
5.	ESPRIT	64%
6.	3 and 4 combined	72%
7.	3 and 4 combined with GMM	80%

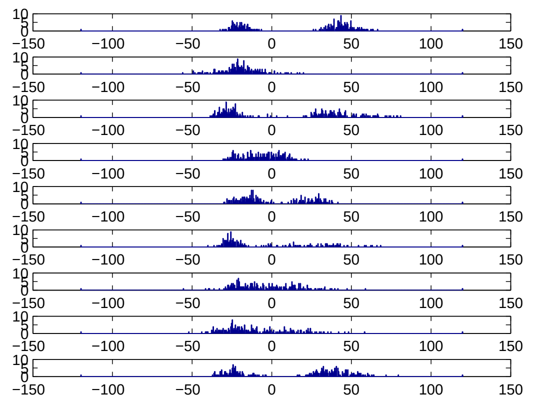

Fig. 10: Within and between class variation of Zero Crossing Feature.

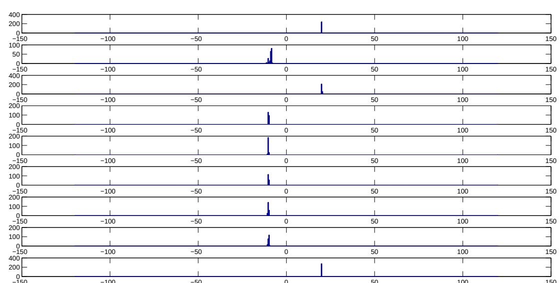

Fig. 11: Within and between class variation of Mean.

A major difficulty in our classification we faced was to correctly classify the signals that didn't belong to any of the classes in our training set. We tried several approaches in this regard. The most trivial method was to threshold the maximum probability at some value. If the highest probability output by the SVM for a signal is less than this value we would classify it as not belonging to any of the classes. However, we found that such a threshold could not be chosen. Any reasonable threshold misclassified several signals as belonging to 'Not any Class'.

E. Other Experimentations

1.It was inferred, insufficient number of samples of audio recordings from each grid given to us could mean a larger misclassification probability for audio recordings. To overcome this, synthetic white noise was added to clean power signal to model the noisy nature of the ENF signals from audio recordings.[ref] Now separate models were trained using the clean and the noise added ENF signals. Their class probabilities were combined as above. By this approach the classification accuracy decreased by 8% for the practice signals.

2.From the probability estimates for each class given by SVM, we could say that about 12% of the signals could have been classified correctly, had the class with second highest probability chosen. For this we implemented a binary SVM between the classes with the two highest probability estimates using a reduced number of features that best discriminate the two classes. We had difficulties in deciding the most discriminating set of features for every pair of classes. Initially we used those features whose absolute difference of mean values for the two classes was highest. Another approach was to select those features whose variance was least in each class. The accuracy achieved using both were similar but less than the original model.

3.We took five random power signals from each grid and added a small fraction of a signal taken randomly from a different grid of the same nominal ENF frequency. Thus we generated five extra signals for each grid. Our intention was to make the classifier more robust to misclassification. But the accuracy dropped by 4% after this method.

F. Results

After combining the models based on Frequency Tracking and Quadratic Interpolation methods, and using GMM to identify the signals not belonging to any class, we got an accuracy of 80% on the practice dataset. The classification result obtained by our final model on the two datasets are given below:

1) Practice Dataset Result: :

AHCFD,DEIND,AFDDC,INNAE,DBBID, CDFGB,DHCHG, ECIHI,EHECF,FNGEI

2) Testing Dataset Result: :

NDDCD,FFEAF,CIGBG,DFCEH,HHHDD, NIDAI,DNGHI,IDCBG,ENIBG,FGIAD, CIAID,HAEEC,IHDDG,CECBI,EICDI, BDBDB,DINAG,IABIH,IIDEA,GBFDD

V. Circuit Design and Data Analysis for ENF Acquisition

In this section we explain the details of the power recording device at our location.We have developed a simple and inexpensive but efficient technique for obtaining the power recordings.

A. About the circuit

First we step-down the supply voltage and then attenuate the same to a the voltage level compatible for the following section of the circuit, Analog to Digital Converter(ADC) . The range of voltage for measuring capable by the ADC is also noted. In our circuit design, a 6-0-6 Volt 500mA Transformer is used. The stepped down voltage has a maximum voltage of 9V which is brought down to a 2.25V through a voltage dividier circuit with high value resistors (150k Ohm and 450k Ohm) to limit the current and thus reduce the power dissipation. Figure 12,13,14 shows schematic diagram of circuit,photograph of the circuit and TRRS interfacing with a computer.

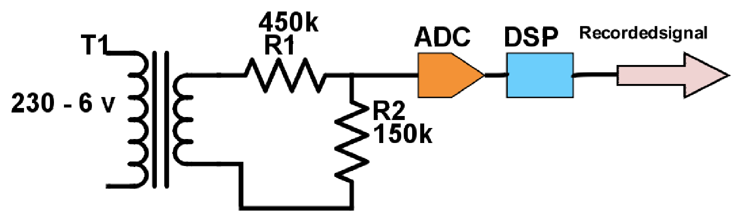

Fig. 12: Schematic Diagram of Circuit.

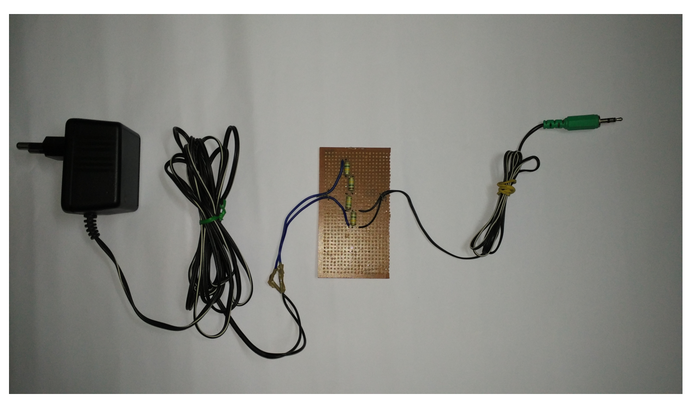

Fig. 13: Circuit used for recording.

Fig. 14: 3.5 mm TRRS jack.

Depending on the sampling rate of the ADC, an anti aliasing filter should also be placed in the circuit along with a fuse for safety purposes. For an ADC we us the inbuilt sigma-delta ADC from the sound card on the computer. The Sigma Delta converters in the sound card which provide inherent anti-aliasing protection. Since the signal of interest has such a low frequency, the sampling rate can be correspondingly low. Here we take sampling rate of 1000Hz and record using Audacity Software as shown in Figure 15.

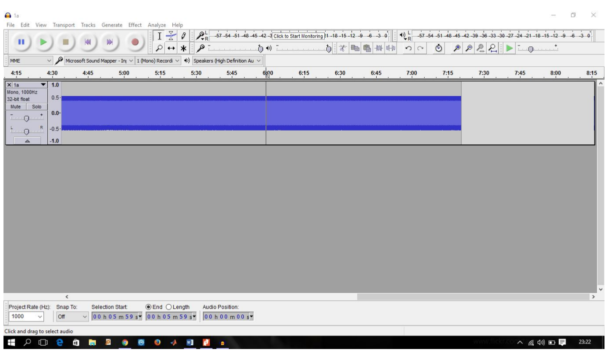

Fig. 15: Audacity software.

B. Sigma-Delta ADC

Sigma-delta A/D converter consists of an oversampling modulator, followed by a digital filter and a decimator. The modulator output swings between two states (high and low), and the average output is proportional to the magnitude of the input signal. Since the modulator output always swings full-scale (1 bit), it will have large quantization errors. The modulator, however, is constructed so as to confine most of the quantization noise to the portion of the spectrum beyond 500Hz.

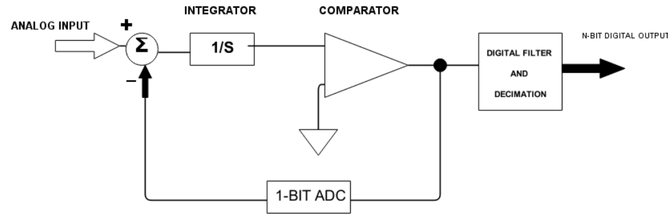

Fig. 16: Schematic diagram of Sigma-Delta ADC.

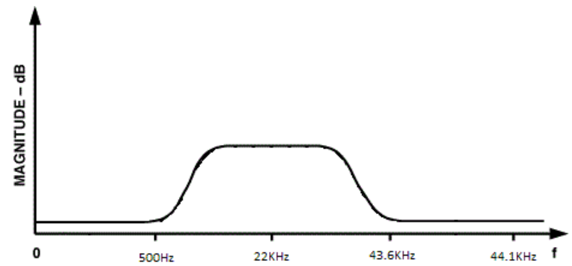

Fig. 17: Output of modulator.

As shown in Figure 17, the ENF signal spectrum at a ranges from 0 to 500Hz and 43.6 KHz to 44.1 KHz l, while the quantization noise is at a region between them, where 44.1 KHz is the sampling rate of the modulator.

The digital filter is an n-tap FIR filter and takes the high-speed low-resolution (1-bit) modulator output and performs a weighted average of n modulator outputs in a manner dictated by the desired filter characteristics. The output of the filter is a high-resolution word, which becomes the A/D output. The digital filter is designed to filter out between 500Hz and 43.6kHz. Cleaning out all the noise in between 500Hz and 43.6kHz makes it possible to reduce the sampling rate to values between 1KHz to 44.1KHz without causing any aliasing.

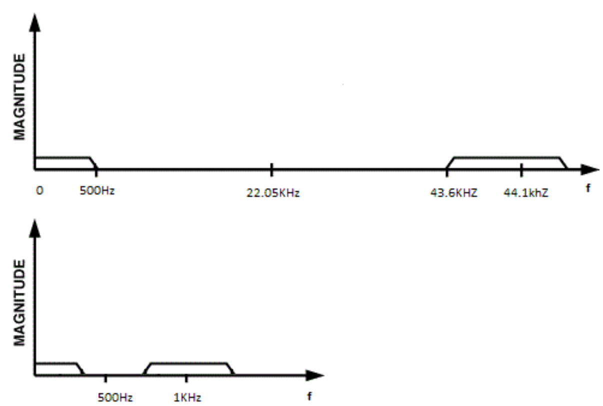

Fig. 18: Output of modulaor after digital filter(top) and decimator(bottom).

The upper figure 18 shows the output of the modulator after digital filtering but prior to decimation. The lower figure shows the spectral output after decimation-the final A/D output.The more detailed description can be found in [20]

C. Analysis of recorded power signals

Using the above setup, we obtained power recording for a duration of 10hrs. To get an overall idea about the variation in ENF with time and load, the signal was recorded at different times from varying locations. The PSD plot of the power recording, shown in the Figure 19, clearly depicts that power exists at the nominal frequency of 50 Hz as well as its harmonics.

We used Quadratic interpolation and Frequency tracking to estimate the ENF signal from the recording and tested the extracted signals on the classifier we developed, the results obtained are tabulated. The signal was judged by the GMM as not belonging to any of the above classes.Table III shows a brief analysis of the recorded Power Signal and how it varies through out a day.

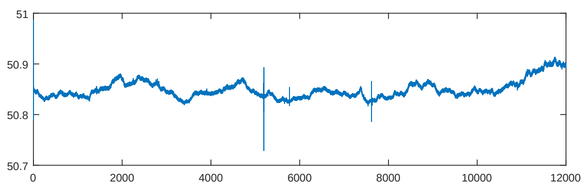

Fig. 19: ENF extracted from our recorded power during early hours using Quadratic Interpolation

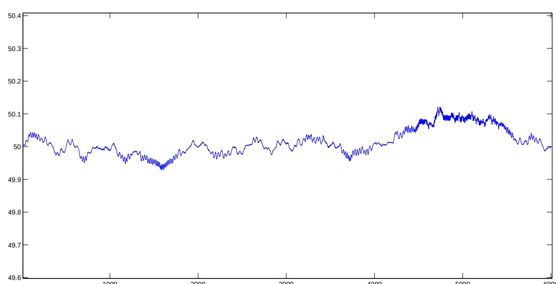

Fig. 20: ENF extracted from our recorded power at noon using Frequency Tracking Algorithm.

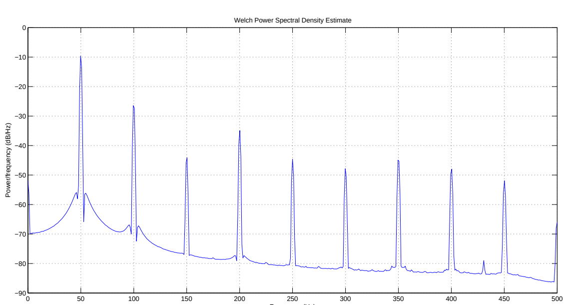

Fig. 21: Power spectral density of our power recording(1-4).

TABLE IV: Analysis of Extracted Signal

Time	Mean	Variance	range	Noise Variance of AR(5)
Morning	50.09	6.1×10^{-4}	0.13	1.87×10^{-4}
Noon	49.97	0.0029	0.26	2.9×10^{-4}
Mid-night	50.01	0.0015	0.19	2.13×10^{-4}

VI. Conclusion

This report described the various techniques we implemented for extracting ENF signal from power as well as audio recordings .The ENF can be classified without having the power or audio references. We used the SVM classifier for classifying the extracted ENF signals. We report a classification accuracy of 80% as verified by the link provided by organising team SP CUP. We have presented the results obtained on the test data set. We also presented the details of the hardware used for power recordings at our location .

References

[1] Ravi Garg, Adi Hajj-Ahmad, and Min Wu, *Geo-Location estimation from electrical network frequency signals*, ICASSP 2013.

[2] Adi Hajj -Ahmad, Ravi Garg, and Min Wu, *ENF Based Location Classification of Sensor Recordings* ICASSP 2013.

[3] Hui Su, Adi Hajj-Ahmad, Min Wu, and Douglas W. Oard, *Exploring the use of ENF for multimedia synchronization* 2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP).

[4] Hui Su, Ravi Garg, Adi Hajj-Ahmad, and Min Wu, *ENF analysis on recaptured audio recordings* ICASSP 2013.

[5] Adi Hajj -Ahmad, Ravi Garg, and Min Wu, *Instantaneous frequency estimation and localization for ENF signals* ICASSP 2013.

[6] Philip Top, Mark R. Bell, Ed Coyle, and Oleg Wasynczuk, *Observing The Power Grid* IEEE Signal Processing Magazine 2012.

[7] Xin ZHOU, Ying WU, Bin YANG *Signal Classification Method Based on Support Vector Machine and High-Order Cumulants* Wireless Sensor Network, 2010, 2, 48-52 doi:10.4236/wsn.2010.21007 published online january 2010

[8] 2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)

[9] Wei-Hong Chuang, Ravi Garg, and Min Wu, *Anti-Forensics and Countermeasures of Electrical Network Frequency Analysis* IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 8, NO. 12, DECEMBER 2013

[10] Ravi Garg, Avinash L. Varna , Min Wu, *Modeling and Analysis of Electric Network Frequency Signal for Timestamp Verification* .

[11] Adi Hajj -Ahmad, Ravi Garg, and Min Wu, *ENF-Based Region-of-Recording Identification for Media Signals* IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 10, NO. 6, JUNE 2015.

[12] A. C. Kot, S . Parthasarathy, D. W . Tufts, and R. J. Vaccaro, *Statistical Performance of Single Sinusoid Frequency Estimation in White Noise Using State-Variable Balancing and Linear Prediction*

[13] Ode O jow u,Johan Karlsson,Jian Li, and Yilu Liu, *ENF Extraction From Digital Recordings Using Adaptive Techniques and Frequency Tracking* IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL.7,NO.4, AUGUST 2012.

[14] Avinash.L.Varna,Adi Hajj -Ahmad, Ravi Garg, and Min Wu, *"Seeing" ENF"Power -Signature-Based Timestamp for Digital Multimedia via Optical Sensing and Signal Processing* IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY,VOL.8,NO.9, SEPTEMBER 2013.

[15] Ling Fu,Penn N.Markham,Richard.W.Conners,Yilu Liu, *An Improved Discrete Fourier Transform-Based Algorithm for Electric Network Frequency Extraction* IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY,VOL.8,NO.7, JULY 2013.

[16] Adi Hajj -Ahmad, Ravi Garg, and Min Wu, *Spectrum Combining for ENF Signal Estimation* IEEE SIGNAL PROCESSING LETTERS,VOL.20,NO.9,SEPTEMBER 2013.

[17] Xin ZHOU, Ying WU, Bin YANG, *Signal Classification Method Based on Support Vector Machine and High-Order Cumulants* ICASSP 2013.

[18] Richard Duda,Peter Hart,David Stork, *Pattern Classification* Wiley Publications.
[19] Shu Lin,Daniel J Costello, *Error Control Coding* Pearson Education India.
[20] J.C. Candy and G.C. Temes, *Oversampling Delta-Sigma Data Converters-Theory, Design, and Simulation* , IEEE Press, Piscataway, NJ, 1991.

About The Zenith

Dr.Jiji CV is Professor and Head of Department of Electronics at College of Engineering Trivandrum.
He had his BTech from TKM College of Engineering, Kollam and MTech and Ph.D from Indian Institute of Technology, Bombay .His major interests are advanced digital signal processing,Multirate systems and wavelets,Digital image processing, Estimation and detection theory.He has published many papers related to Image Super resolution and Image processing.

Vineeth Warrier, Hari Govind V K, Sreeshan M S,Vishnu K P are final year students , pursuing their B-Tech degree in Electronics and Communications Engineering at College of Engineering, Trivadrum

Unnikrishnan S Nath, Noufal Puthukkudi, Vishnuprasad K, Gowri S Gopinath, Anjana Raju, Lakshmi Jayalal are pre-final year students, pursuing their B-Tech degree in Electronics and Communications Engineering at College of Engineering, Trivadrum.

Sarath Sasidharan is currently doing his post graduate studies in the Electronics Department of College of Engineering, Trivandrum