ENF Based Grid Classification System: Identifying the Region of Origin of Digital Recordings

Riyasat Ohib, *Student Member, IEEE,* Samin Yeasar Arnob, *Student Member, IEEE,* Riazul Arefin, *Student,* Mohammad Muhtady Muhaisin, *Student Member, IEEE,* Tanzil Bin Hassan, *Student,* MR Amin, Md. Taslim Reza, *Student Member, IEEE*

***Abstract*— The Electric Network Frequency (ENF) is the supply frequency of power distribution networks which can be captured by multimedia signals recorded near electrical activities. It normally fluctuates slightly over time from its nominal value, which is usually of 50 Hz/60 Hz. The ENF remain consistent across the entire power grid. This has led to the emergence of multiple forensic application like estimating the recording location and validating the time of recording. In this report we examine an ENF based Machine Learning system which infers the power grid in which the multimedia signal was recorded. We worked on different features which serve as signature for power grid. Then we used those in a multiclass machine learning implementation which is able to identify the grid of the recorded signal.**

Keywords— Electric Network Frequency, grid, forensics, estimation, machine learning.

I. INTRODUCTION

The supply frequency of the power distribution network i.e. ENF has a nominal value of 60/50 Hz. The instantaneous frequency fluctuates about the nominal value due to the load control mechanisms and the changes in the load demands within the power grid. The fluctuations of the ENF although random, are unique within a particular electrical network [1]. The tendency of these variations, at a particular time, are almost same throughout the same grid. These variations of the ENF over time is defined as the ENF signal.

The detection of tampered and forged audio/video recording is of immense importance for media forensics. Prior to the increased use of digital recorder, forensic audio/video analysis relied on different techniques for analog recorders. But for digital recordings, altercations can be made very easily without leaving behind any trails, because digital recorders produce a recording by converting sound variations to a series of numbers [2] making authentication of these recordings a lot more difficult. As a result the use of ENF signals to identify tampered or modified audio recordings was proposed by Grigoras in 'Electric Network Frequency (ENF) Criterion' [3], [4], [5].

The ENF signal is presently being used in multimedia forensics applications, as it gets embedded in the multimedia recordings made in the vicinity of electrical activity. The audio recordings can pick up the signals due to the mechanical or acoustic hums or electromagnetic interferences from the power lines. The clean power recordings can be extracted using an audio recorder connected directly to the power mains via a step-down transformer [14]. Applying a Band-Pass Filter around the nominal frequency and employing a frequency estimation algorithm, the dominant frequency surrounding the nominal frequency can be estimated frame-by-frame, thus forming the ENF signal.

Analysis based on ENF signals can also approximate the time of recording of audio and video files or detect tampering or modification in those files [5], [7], as ENF fluctuations extracted from clean power signal is same as the embedded ENF variations found in multimedia recordings [5], [6]. These applications generally require prior information of the grid of origin or simultaneous power references to identify the grid of origin. Moreover they are also computationally inefficient as a large archive of power reference is required for them to be of use. To overcome this problem a Machine Learning approach was proposed in [12].
The ENF signals from a particular grid although random, vary from those of other grids in their nature and manner of variations. Statistical

features extracted from the ENF signals of different grids can be used to develop a Machine Learning system that can identify the grid of origin, and therefore the region of recordings.

We were provided with power and audio recordings from nine different grids around the world, which includes both 'Noise-Free' power recordings and 'noisy' audio recordings. Using statistical features derived from the ENF signals, we tried to differentiate the region of recording of the media signals based on a Machine Learning approach.

The ENF signals picked up in these recordings can be readily extracted directly from the Electric mains using a step down transformer and a simple voltage divider circuit.

II. ENF EXTRACTION

In this section we describe the database of Power and Audio recordings provided to us. Subsequently, we discuss the procedure and methodology adopted to extract the ENF signal from the recordings. We end this section by analyzing the similarities and dissimilarities between the extracted ENF signals, and discussing the statistical features that might be used as discerning features for the classification of the ENF signals.

A. Database Description

Recordings from nine different grids titled as A, B, C... to I were provided to us for analysis in the IEEE Signal Processing Cup 2016. Among them we found the nominal frequencies of grids A, C and I to be 60 Hz and the rest of the grids to be 50 Hz through Fourier Transform Analysis. The nominal frequencies corresponding to each grids and their range of variation are provided in Table-I. All the grids contained both power and audio recordings among which the duration of power recordings were of ten to sixty minutes, while all the audio recordings were of thirty minutes each. The Database consisted of variable number of power recordings but same number of audio recordings across all grids. For each power and audio files, we at first divided the main recording into segments of ten minutes each. Then, among these segments, we divided the signals into non-overlapping frames of 5 seconds duration.

Table I
DESCRIPTION OF THE DATABASE

GRID	Nominal frequency (Hz)	Maximum frequency Hz	Minimum frequency Hz	Frequency Range Hz
A	60	60.032	59.651	0.067
B	50	50.850	49.142	1.707
C	60	60.044	59.962	0.082
D	50	50.057	49.942	0.115
E	50	50.077	49.925	0.152
F	50	50.778	49.942	0.134
G	50	50.142	49.801	0.341
H	50	50.316	49.716	0.600
I	60	60.047	59.951	0.095

The Spectrum Combining approach [8] is used to estimate the instantaneous ENF frequency for a time frame. This approach utilizes the presence of the ENF components in the nominal base frequency and in multiple harmonic bands to arrive at a more precise estimate. We used the Spectrum Combining method in conjunction with the Short Time Fourier Transform (STFT) [9], [11] for estimating the dominant frequency component of each time frame, thus generating the ENF signal.

B. Spectral Combination and SNR Calculation

The Spectral Combining procedure suggests producing an ENF estimate by combining base and multiple harmonic spectral bands, each weighted according to its relative strength with respect to noise. In this sub-section we will recall the method of Spectral Combination, along with the derivations and equations from [8]. The estimation is made on a frame-by-frame basis using STFT over time, where the deviation Δf_0 of the base ENF from the nominal frequency f_0 is approximated. For a given STFT frame, the observed power spectrum component, $P_{B,\lambda}(f)$ contributed at the harmonic band λ, can be expressed as

$$P_{B,\lambda}(f) = A_\lambda h_\lambda(f) + \mathrm{P}_{n,\lambda}(f) \qquad (1)$$

For $f \in [\lambda(f_0 - f_B), \lambda(f_0 + f_B)]$,

Where f_B reflects the frequency band of ENF presence around the base frequency, A_λ is the magnitude of the energy of the frequency component surrounding λf_0 and $\mathrm{P}_{n,\lambda}(f)$ describes the white noise component around the same harmonic within the bandwidth. The

$h_\lambda(f)$ in equation (1) is an impulse-like function which achieves its maximum value at $f = (f_0 + \lambda\Delta f_0)$. Compressing and shifting the spectrum components to the base range of $[f_0 - f_B, f_0 + f_B]$ and then combining the components together, the weighted summation can be represented as

$$S(f) = \sum_{\lambda=1}^{L} w_\lambda P_{B,\lambda}(\lambda f) \qquad (2)$$

The combining weight w_λ, in equation (2) weighs each harmonic spectral bands according to their signal-to-noise ratio (SNR). The dominant frequency $f_{ENF} = f_0 + \Delta f_0$, representing the ENF estimation of that time frame can then be calculated by taking the maximum value in the $S(f)$ vector.

We computed the combining weights over a ten minutes duration of recording and computed it again at the start of the next ten minute segment to make the weight adaptive to the changing strength of the ENF in the bands over time. For example, the combining weight of the first harmonic w_1 can be represented by $P_{signal,1} / P_{noise,1}$. We analyzed the given data set and determined the range of fluctuations for different grids in Table I.

We observed that, the ENF fluctuations of the 50Hz grids were comparatively higher than those of the 60Hz grids. We measured the ENF range of 60Hz and 50Hz grids and chose the maximum range belonging to each of the two frequency groups. Then we empirically set f_B =1Hz, for the 60Hz grids and f_B =3.2Hz for the 50Hz grids such that the ENF bandwidth of all the grids might fall into our chosen bandwidth. The overall band for computing SNR would be [59, 61] Hz for 60Hz electrical networks and [46.8, 53.2] Hz for 50Hz electrical network. Then $P_{signal,1}$ is the average Power Spectral Density (PSD) within the band [59.95, 60.05] Hz and [49.14, 50.85] Hz over the chosen time duration and $P_{noise,1}$ is the average PSD in the bands [59, 59.95] Hz and [60.04, 61] Hz, and [46.8, 49.14] Hz and [50.85, 53.2] Hz for the 60Hz and 50 Hz grids respectively over the same time duration. For the ease of computation, we compressed and shifted the higher harmonic spectral bands to the nominal $P_{signal,1}$ and $P_{noise,1}$ bands and computed their weights.

After calculating the weights we perform STFT, where within each time-frame we employ the Spectral Combination approach just discussed, making use of the weights calculated in this sub-section to find out the frequency estimation for each time frame.

C. Filtering of Harmonic Components

For implementing the Spectral Combining approach, we filtered out the harmonic components $\lambda = 1, 2, 3...Q$, (where Q = 8) using a bandpass IIR elliptic filter of the order 40. Elliptic filters were chosen empirically to provide sharp cutoff frequency and customized Stop-Band and Pass-Band ripples. The filtered harmonic components were then used as time domain recording data for frequency estimation utilizing the Spectral Combining approach using the STFT method.

D. ENF Estimation of Time Frame by STFT

The frequency estimation process to be used must provide adequate frequency resolution and sampling interval. The method we chose to track the dynamic behavior of the ENF is based on the Short Time Fourier Transform (STFT) [9]. The estimation of the peak magnitude was achieved using a quadratic interpolation technique with a mild zero padding scheme [10]. We summarize the technique that we followed, reported in [11] along with necessary equations and derivations.

The STFT was implemented by dividing the original signal data sequence $x[n]$, $0 \le n \le N-1$, into J segments of length M samples:

$$x_m[n] = x[mL + n],\ 0 \le n \le M-1 \qquad (3)$$

Where x_m is the m^{th} frame of the input signal and L is the 'hop size' or the number of samples advanced between each time frames. Each frame is then multiplied by a length M spectral analysis weighting window w producing:

$$x_m[n] = x_m[n].w[n] \qquad (4)$$

We chose a 'Hanning Window' with a window size M of 5000 samples. For increasing the frequency resolution the data from the windowed frame is then extended by zeroes by a zero padded windowed frame $x_m[n]$.

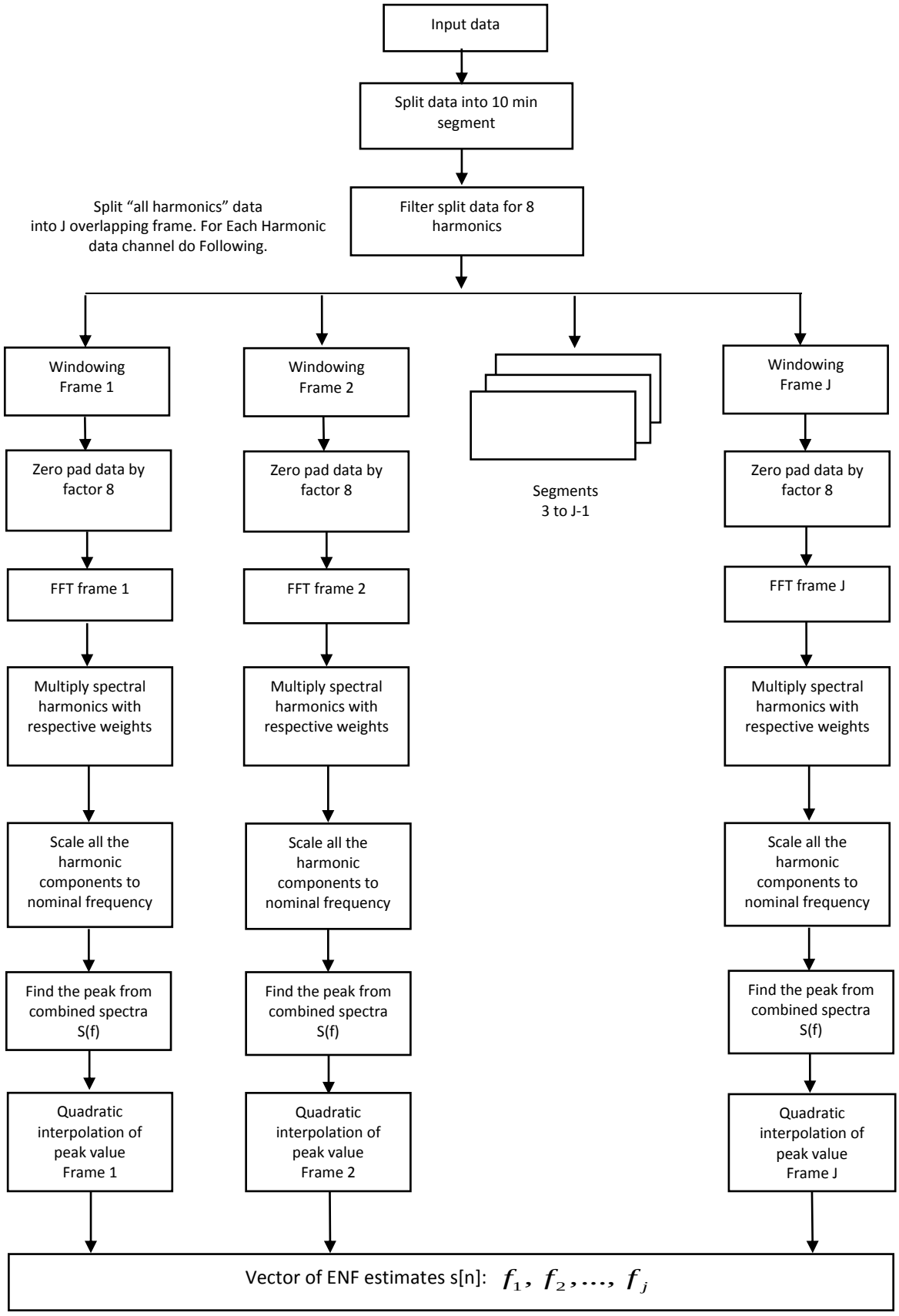
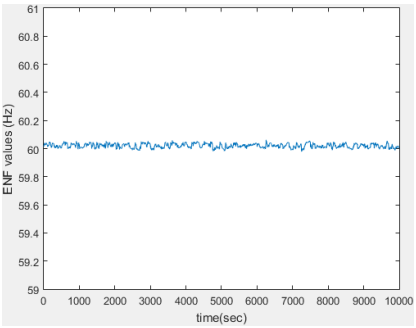
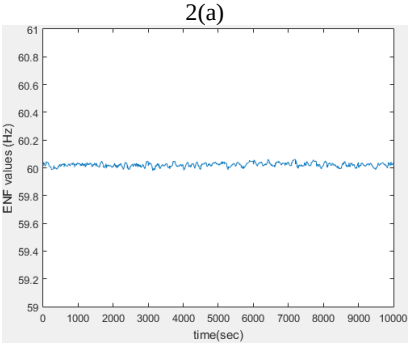


Figure 1. Program flow chart



2(a)



2(b)

Fig 2: ENF signals extracted from 60Hz Grids (a) Grid A. (b) Grid C. (c) Grid I.

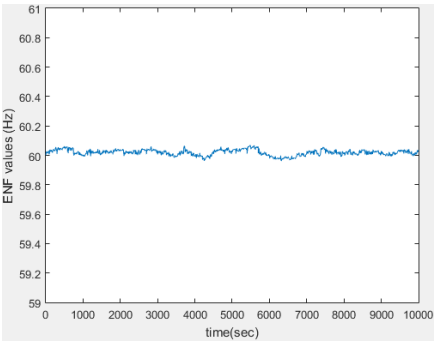
Each frame is then converted to the frequency domain using a length P FFT to produce the STFT at frame m as:

$$X_m[k] = \frac{1}{P}\sum_{n=0}^{P-1} x_m'[n] e^{\frac{-j2\pi kn}{P}} \qquad (5)$$

Where is k is the k^{th} frequency bin. We set the hop size L to 5000 data samples to match the window length M and create a non-overlapping STFT scheme. The STFT segmentation process is elucidated by the Figure-1.

It is unlikely that, the ENF peak value in any frame will fall exactly into a frequency bin of an FFT Transform point. By zero padding in the time domain, the FFT size is increased which in turn makes the FFT bin bandwidth f_s / p proportionally narrower. The result is a more accurate interpolation in the frequency domain and a more densely sample spectrum. Conversely, for attaining reasonable accuracy, the zero padding factor has to be quite large resulting in a high FFT size. To maintain a balance between computational efficiency and accuracy we used a zero padding factor b of 8.

We performed the STFT on all the Q harmonic channels simultaneously to find the dominant



2(c)

frequency surrounding each harmonics. While estimating the frequency deviation Δf_0, we first compressed and shifted the spectral components of the higher harmonics to the nominal base range $[f_0 - f_B, f_0 + f_B]$ and combined the components together according to equation (2) to get $S(f)$, which denotes the weighted summation of the harmonic spectral components. The combining weights w_k as calculated in *Section-II B*, were used to weigh the spectral components. The ENF frequency estimation $f_{ENF} = f_0 + \Delta f_0$ was thus obtained by searching for the maximum in $S(f)$. For obtaining a high accuracy, we resorted to a quadratic interpolation scheme in a conjunction with a low zero padding as discussed by Abe & Smith [9]. Using the left and right adjacent bins the spectral bandwidth, we employ the quadratic model

$$v = \frac{1}{2}.\frac{\alpha - \lambda}{\alpha - 2\beta + \lambda} \qquad (7)$$

Where

$$\alpha = 20\log_{10}|X_m(k_{\beta-1})| \qquad (8)$$

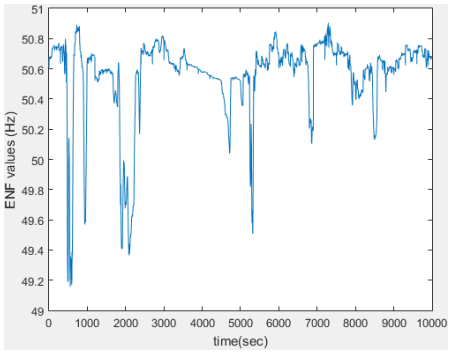
$$\beta = 20\log_{10}|X_m(k_{\beta})| \qquad (9)$$

$$\lambda = 20\log_{10}|X_m(k_{\beta+1})| \qquad (10)$$

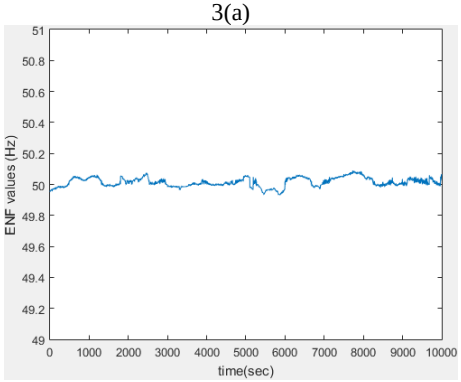
to gain a finer estimate of the ENF frequency for a particular time frame. Employing the same approach, we estimated the ENF frequency for subsequent time frames of a data segment of 10 minutes. Thus generating an ENF signal segment of length $S = 120$ samples.

E. Comparison of ENF Signals From Different Grids

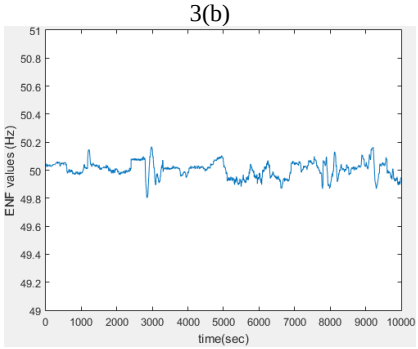
We differentiate here the statistical differences between the ENF signals extracted from different grids. These differences help us to adopt a set of features for our Machine Learning



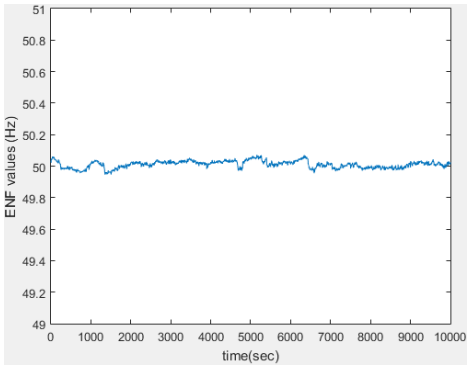
3(a)



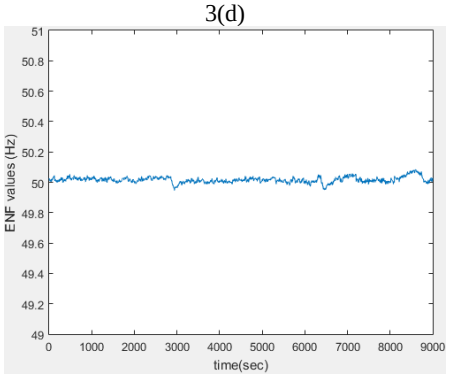
3(b)



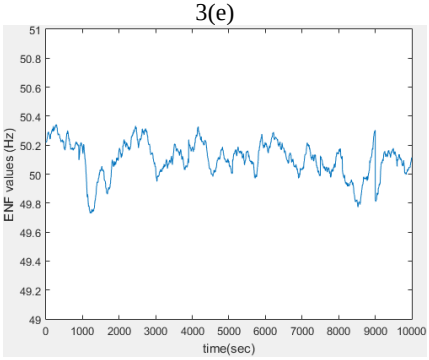
3(c)



3(d)



3(e)



3(f)

Fig 3: ENF signals extracted from 50Hz Grids. (a) Grid B. (b) Grid D. (c) Grid E. (d) Grid F. (e) Grid G. (f) Grid H.

system. In our database three grids have a nominal frequency of 60Hz and six grids have a nominal frequency of 50Hz as mentioned in *Table-I*. The ENF signals extracted from the power signals can be seen in the Figures 1 and 2. The first difference that is evident is the mean of the ENF signals, which can be easily identified depending on how close their average is to the 50Hz or 60Hz mark. Even among the grids with the same nominal frequency, other distinguishable features can also be observed. For example the mean of grid B remains well above 50Hz for most of its duration, while those of grid F remains close to the 50Hz mark.

Besides differing in their means, the ENF signals also demonstrate uniqueness in the pattern of their variations. Scrutinizing the 60Hz grids, titled A, C and I, we found that, they are quite similar in their variations as evident from Figure 1. This suggests a better control mechanisms in these grids. Among these three, only grid 'I' showcases a small difference in its fluctuations, which tends to be a bit slower than those of grids 'A' and 'C'.

Besides differing in their means, the ENF signals also demonstrate uniqueness in the pattern of their variations. Scrutinizing the 60Hz grids, titled A, C and I, we found that, they are quite similar in their variations as evident from Figure 1. This suggests a better control mechanisms in these grids. Among these three, only grid 'I' showcases a small difference in its fluctuations, which tends to be a bit slower than those of grids 'A' and 'C'. In case of the 50Hz grids more notable differences can be observed. The ENF signals of Grid B, for example, demonstrates the most abrupt changes among all the grids as

apparent from Figure 2. This suggests high load variations and poor control mechanism present in the grid.

Similarly, in our analysis, the ENF signals of grids G and H also display more fluctuations than the ENF variations of Grids D, E and F, but these are still not as sharp as those of grid B. Even among the D, E and F grids, some discerning characteristics can be seen. The F grid, for instance exhibits a more controlled variation, whereas grid D appears to drift more before returning to its base frequency range.

III. Location Classification System

Exploiting the unique fluctuations of the ENF signals, a multiclass classification system can be devised to identify the grid of origin of the media recordings. In this section we discuss our implementation of the Machine Learning system based on the work in [11]. We start by discussing the feature extractor part of the Learning system and explain the features used for our System.

We also explore some additional new features from the ones described in [11], which have the potential to aid the Learning System. The second element of our classification system is a multiclass classifier based on the MATLAB's Error-Correcting Output Code (ECOC) multiclass model for Support Vector Machines (SVM). We then discuss the methodology used to train the system and conclude this section by examining the results obtained.

A. Feature Extraction & Analysis

From the empirical differences in the variations of the ENF signals discussed in Section II-E, we can extract meaningful statistical features for our classification system. We took a set of ENF signal segments $s[n]'s$ of fixed size $S = 120$ samples among the power Grids of our dataset. Each of the ENF signal segments was extracted from an ENF-containing record of ten minutes long, as mentioned in section II. In our data-set the number of grids with 50Hz ENF was six and with 60Hz ENF was three. Basing on the system proposed in [11], we adopted the mean and the variance of the ENF segment as well as its dynamic range. Moreover, we applied a wavelet signal analysis to study the ENF signal segments at more than one time-frequency resolutions.

TABLE III:
USED FEATURE COMPONENTS

1	Mean of ENF segment
2	Log(variance) of ENF segment
3	Log(range) of ENF segment
4	Log (variance) of approximation after F-level packet wavelet analysis (F=3)
5-34	Log(variance) of F level of detail signals computed through F-level packet wavelet analysis (F=3)
34-37	AR(4) model parameters
38	Log($e[n]$) after AR modeling
39-42	Polynomial(4) Model Parameters
43-44	Skewness and Log(Kurtosis) of ENF signal.
45-46	Log(co-efficient) of Linear Regression
47-55	Normalized Amplitude of the Q Harmonic components.

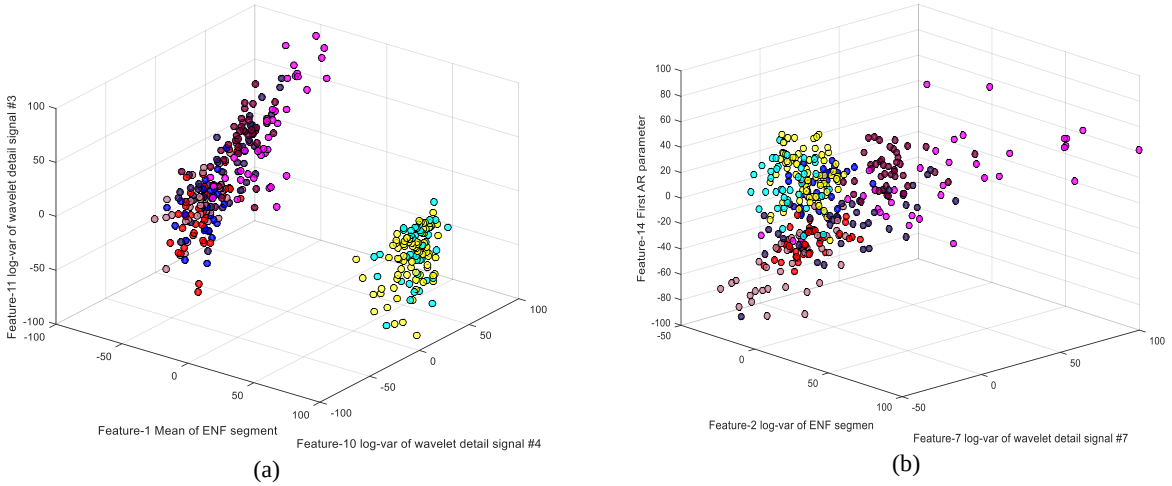
We applied a F level wavelet decomposition where each level provides an approximation to the original signal and the variations at the respective level of resolution [13]. We calculated the variances of the high- pass band of each decomposition level (the details) and also the variance of the lowest time-frequency band (the approximation) as candidate features. The features collected from wavelet based analysis help to capture the fine differences of the ENF signals of distinct electrical networks.

To compliment these features, we also collected a set of features obtained from an auto regressive statistical modeling of the ENF signal. We make use of the model described in [14]. An ENF segment $s[n]$ would be modeled as:

$$s[n] = \sum_{i=1}^{p} a[i]s[n-i] + e[n] \qquad (11)$$

Where $a[i]$ is the i^{th} coefficient of the AR model, $e[n]$ is a white noise with mean zero and p is AR order. We select a fourth order AR model.

For use in the Classification System, five features from this AR modeling: the four parameters of the AR model $a_1, a_2, a_3,$ and a_4, and $e[n]$ are selected. The AR parameters



(a) (b)

Figure 4. Selected normalized feature values of Training data. (a) Features 1,10 and 11 (b) Features 2,4 and 14.

indicate the linear relationship of the samples of $x[n]$ and the $e[n]$ describes how well the signal fits an auto-regressive model.

To focus on the orders of magnitude and to increase the chances of separation between final feature values, we used the log operator on the mean and variance values. Moreover, the features extracted from similar grids form clusters, as evident from Figure 4.

Exploring some additional features to aid our Machine Learning System, we utilized the Fourier Transform profile of the recorded signals. We observe that the magnitude of the harmonic components of the power recordings differ from one grid to another in their magnitudes. To utilize this property we fit a polynomial model over the normalized magnitudes of the harmonic components and use the coefficients and the constant term of the fitted polynomial as features. We also utilized the normalized magnitude of the harmonic components as potential features.

Furthermore, we added two more probabilistic features: skewness and kurtosis of the ENF variations. Skewness is a measure of the asymmetry of the probability distribution of real-valued random variable about its mean. Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution. Moreover, we also applied a linear regression model on the extracted ENF signal and used its co-efficients as features. The features used are summarized in Table-III.

B. Normalization of Features

For use in the Machine Learning System the computed values of the features should be normalized in a linear scale. The range of normalization used in our work is [-100,100], where the c^{th} feature value in an example of the training dataset is normalized with respect to all the other feature values in the position k of the N training examples. While training the Classification System we stored the normalization parameters, and later used them to normalize the Practice and Testing Data provided for analysis in the competition.

We resorted to a similar technique as the one reported in [11]. Equations(13)-(15) stated below describes the process of normalization for features $c = 1, 2, ..., 16$.

$$\mu_c = \frac{1}{M}\sum_{j=1}^{M}\left(\frac{1}{N_j}\sum_{i,l_i=j} f_i[c]\right) \quad (13)$$

$$f_i'[c] = f_i[c] - \mu_c \quad (14)$$

$$f_i[c]'' = 100 \times \frac{f_i'[c]}{\max_i |f_i'[c]|} \quad (15)$$

Here M represents the number of classes, where each class was composed of N_j examples, with $j = 1, 2, ...M$ and $\sum_{j=1}^{M} N_j = N$. The label and feature vector of an example i is denoted by l_i and f_i, respectively, for $i = 1, 2, ...N$, and output the final normalized result f_i''.

C. Multiclass SVM Classifier

In our implementation of the multiclass Support Vector Machine (SVM) to build the location classifier we used the MATLAB's ECOC model. The ECOC model reduces the problem of three or more classes to a set of binary classifiers. ECOC classification requires a coding design, which determines the classes that the binary learners train on, and a decoding scheme, which determines how the results (predictions) of the binary classifiers are aggregated. For example, in case of a three class system, with a 'one-versus-one' coding design, a decoding scheme, and SVM learners, the ECOC model follows the next steps to build the classification model.

Learner 1, in Table IV, trains on examples from Class 1 and Class 2, and treats Class 1 as the positive class and Class 2 as the negative class. The rest of the learners are trained in a similar fashion. Let ρ be the coding design matrix with elements, $m_{\psi l}$ and s_l be the predicted classification score for the positive class of learner l.

$$\psi = \arg\min_{\psi} \frac{\sum_{l=1}^{L} |m_{\psi l}| \, g(m_{\psi l} s_l)}{\sum_{l=1}^{L} |m_{\psi l}|} \qquad (16)$$

So for a total of M classes, the system trains ${}^{M}C_{2}$ binary classifiers; each binary classifier is trained on one of the ${}^{M}C_{2}$ possible pairs of classes, learning to differentiate between the respective two classes. For testing an example, the MATLAB ECOC implementation also provides M probability (confidence) values, where the j^{th} probability value gives the probability that the testing example belongs to the j^{th} class. The data that was provided to us were also imbalanced: we have more power recordings from some grids versus others, which can create over-fitting or biased problems when testing the system. For this reason we used the weighted multiclass ECOC model which automatically adjusts for the imbalance in the data. In our implementation, we used an eleventh order Gaussian Kernel for our SVM.

D. Systems Trained

The data at our disposal was of two types; Clean ENF signals extracted directly from electrical mains, and noisy ENF signals extracted from the audio recordings. The power recordings generally had ENF segments with high Signal to Noise Ratio (SNR), while those extracted from the audio recordings were quite noisy. For this reason, we opted to train the Machine Learning Systems with two different sets of training data to predict the audio and power recordings. The system trained to predict new power recordings was trained with ENF signals extracted from those power recordings only. On the other hand, as the number of audio files provided for training the system was less in comparison to the power files, so we decided to train the Learning System for classifying audio files on both the Power and the Audio training recordings. This ensured that we had sufficient feature points to predict future data while not compromising on the accuracy of our system.

TABLE IV:
ONE-VERSUS-ONE CODING DESIGN

	Learner 1	Learner 2	Learner 3
Class 1	1	1	0
Class 2	-1	0	1
Class 3	0	-1	-1

For testing the Practice and Test recordings, provided in the competition for classification, we separated the Audio and Power recordings for testing them on the two separate Systems. We achieved this by noting the presence of dominant frequency in the even harmonics of the audio recordings and the odd harmonics in the power recordings. It is also possible to separate the Audio and Power Files on basis of their SNR, but we chose not to pursue this procedure because of its computational inefficiency.

E. Results and Discussion

In this sub-section we present the results obtained from our analysis and justify the methods employed. In the previous section we discussed our approach of developing separate classification systems for audio and power recordings.

For the classification system trained only on power recording we achieved an accuracy of

TABLE V:
CONFUSSION MATRIX

Testing Classes	No. of Examples	A	B	C	D	E	F	G	H	I
A	36	82.72								
B	38		99.38							
C	41			82.67						
D	42				83.62					
E	40					97.62				
F	30						95.21			
G	42							88.62		
H	41								98.88	
I	42									83.46

90.01% with 5 fold Cross-Validation scheme. For this calculation, we used the MATLAB's Machine Learning and Statistics Toolbox functions. The cross validation scheme selects a number of folds (or divisions) to partition the data into. Each fold is held out in turn for testing. We then trained a model for each fold using all the data outside of that fold. We tested each model performance using the data inside the fold, then calculated the average test error over all folds. This method gives a good estimate of predictive accuracy of the final model trained with all the data present in the whole training dataset. Table V lists the confusion matrix thus generated.

On the other hand for the classification system trained on both the power and audio recordings, using the same cross validation scheme we achieved an accuracy of 84.25%. We passed all the practice data through the multiclass learner developed using the above mentioned two models and identified the grid of origin of the 50 practice data provided for our analysis as follows:

AHCFF,EGCND,AFBDC,INNAE,FBBAD,CGNGB,DDCHE,EAIHI,EHECF,FNNEC.

Using the feedback system of the IEEESP Cup webpage, our model predicted accurately 88% of the practice data.

A Similar multiclass classification system was applied to the 100 testing data provided for the first part of the IEEE Signal Processing Cup. The Predicted Grid Label for the testing dataset is listed below:

NDDCE,NHDAF,IEGBG,BDNEH,GNHHG,HFGAI,DNFHN,IECBE,ENNBE,NGEAG,IINNG,HAENC,CCFDG,CEIGI,ENCEE,BEBHA,GNNCG,AABAH,CNDBA,GBFBB

F. Explanation of Developed Program

For developing the ENF extraction and Classification program, we used MATLAB R2015a. We used a combination of scripts and functions in our software. There are four main components in our program: The Gatherer, the Splitter, the Processor and the Classifier. The gatherer script collects the data from each power and audio recordings one by one and sends the recordings to the Splitter. In this portion of the program, a whole audio file is split into ten minute segments of data and sent into the 'Processor' section of the program where the main ENF extraction algorithm is implemented. In the Processor, the Q harmonic bands of the recorded signal is filtered out for Spectral Combination and STFT as described in Section-II-B & D. Thus the ENF signal is generated as a MATLAB vector, and from this ENF signal the features described in Section-III-A are extracted.

The next part of our program is the Machine Learning system where the Multiclass SVM is implemented. The extracted features from the ENF signals of all the training data are normalized and a classification system is trained.

IV. HARDWARE IMPLEMENTATION

A circuit was built to collect the power recordings from the Bangladeshi power grid for

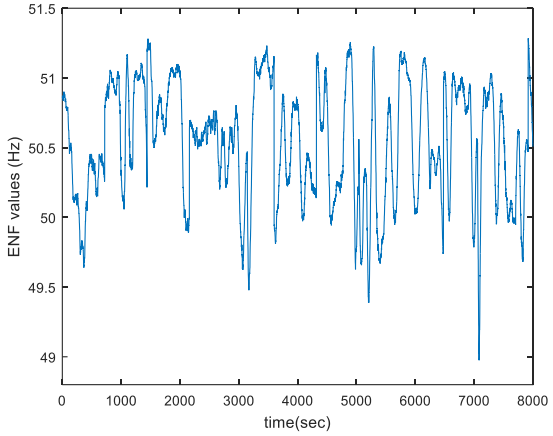


Figure 5. ENF variation of collected signal through our sensing hardware.

analyzing the ENF variations present. The setup was built with least cost possible while not compromising the efficiency and accuracy. After completing the circuit, more than ten hours of reference power recordings were collected at different times of the day in different days of the week. With the help of a voltage divider, 6V was converted into 200mv (p-p). We avoided the use of passive circuit elements while building the setup, so that unnecessary frequency fluctuation can be avoided. This setup was connected with the sound card of a Computer. Thus the originally transmitted electrical signal was recorded with an audio recorder software for ENF extraction using our developed program.

A. Methodology

To obtain the power data of the grid, 220 V was stepped down to 6 V with a transformer. Then, to bring the voltage level down to the low acceptable range of the sound card of the computer, we resorted to a simple voltage divider circuit. The 6 V transformer output was converted into 200 mV (p-p), and connected to the soundcard of a computer. Bearing in mind the change of load conditions at different times of the day, we collected power recordings during mornings, afternoon, midnight and early mornings of different days of the week.

B. Components Used

- One 220V-6V step-down transformer
- Resistors of value 100kΩ & 10kΩ
- One 3.5mm headphone jack
- Jumpers
- Sound Card of a laptop

C. Schematic Diagram

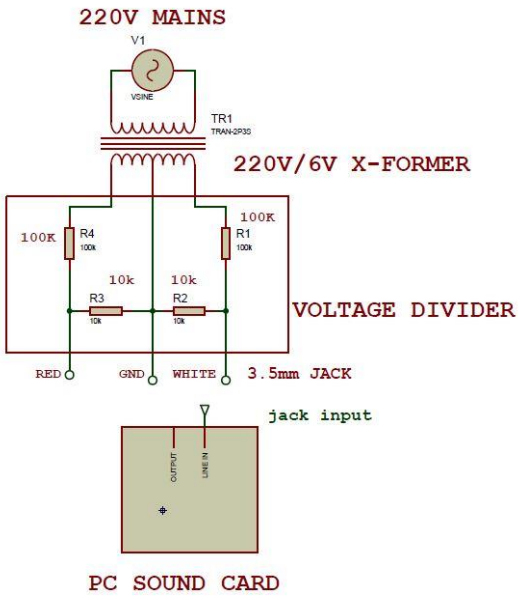


Figure. 6

D. Analysis of Recorded ENF Signal

We collected more than ten hours of power recordings from the Electric mains from Gazipur District of Bangladesh. After extracting the ENF signals from the power recordings using our program, we noticed some similarities and dissimilarities of the signal from the ENF signals of the power recordings of Grids provided to us. We immediately notice the high dynamic range of the ENF signal collected from Bangladesh in Figure 7. The ENF signal fluctuated from around 49.5Hz to over 51Hz for most of its duration. Among the ENF signals of provided grids, only grid B shows some form of similarity with the ENF variations from our collected recordings. The dynamic range of Grid B is also quite large, but Grid B exhibits abrupt and sharp changes in its ENF variations whereas the Bangladeshi Grid displays consistently high variations. This might be because of the poor control mechanism and high load variation in the National Grid of

Bangladesh, which is becoming increasingly obsolete.

To further observe the difference of the collected ENF signals with those from the provided training data we trained our classification system with the both the collected power recordings and the training ones. The resulting confusion matrix presents how well our classification system was able to predict the Bangladeshi grid from its extracted ENF signals.

V. CONCLUSIONS

In this work, we described and implemented a Machine Learning System to identify the Grids of origin of an ENF signal and explored some statistical features that might aid the Machine Learning system's ability to identify the grids of origin. We presented our implementation of an ENF extraction Algorithm and also discussed a Classification system best able to predict Power and Audio recordings. The ENF signals of different grids exhibit different statistical characteristics. Usually this happens due to the size and the quality of the control mechanisms of the grid along with the load variations present in that Electric Network. We have presented a discussion on the similarities and dissimilarities between the ENF signals of different grids provided in the training dataset in this paper. Moreover, we presented our predictions on the 'Practice' and 'Testing' datasets provided for the IEEE Signal Processing Cup 2016. On the Practice Datasets, we achieved an accuracy of 88%, using our program.

In addition we have also explored some new features that we believe might enhance the capabilities of the Machine Learning system to identify the region-of-recordings further.

REFERENCES

[1] M. H. Bollen and I. Gu, "Signal Processing of Power Quality Disturbances." Hoboken, NJ, USA: Wiley, 2006.

[2] R. C. Maher, "Audio forensic examination-Authenticity, Enhancement, and Interpretation," IEEE Signal Processing Mag. , vol. 26, pp.84-94, Mar. 2009.

[3] C. Grigoras, "Digital audio recording analysis–the electric network frequency criterion", *The International Journal of Speech Language and the Law*, vol. 12, no. 1, pp. 63-76 (2005).

[4] C. Grigoras, "Applications of the ENF Criterion in Forensic Audio, Video, Computer and Telecommunication Analysis", *Forensic Science International*, vol. 167, pp. 136-143 (2007).

[5] C. Grigoras, "Application of ENF Analysis Method in Authentication of Digital Audio and Video Recordings", 123rd Audio Engineering Society Convention, New York, USA, (2007), Convention Paper 1273.

[6] A. Hajj-Ahmad, R. Garg, and M. Wu, "Instantaneous frequency estimation and localization for ENF signals," *in Proc. APSIPA Annu. Summit Conf.*, Los Angeles, CA, USA, Dec. 2012, pp. 1–10

[7] R. Garg, A. L. Varna, and M. Wu, "'*Seeing' ENF: Natural time stamp for digital video via optical sensing and signal processing,"* in Proc. 19th ACM Int. Conf. Multimedia, Scottsdale, AZ, USA, 2011, pp. 23–32.

[8] A. Hajj-Ahmad, R. Garg, and M. Wu, "Spectrum combining for ENF signal estimation," *IEEE Signal Process. Lett.*, vol. 20, no. 9, pp. 885–888, Sep. 2013.

[9] J. B. Allen, "Short Term Spectral Analysis, Synthesis, and Modification by Discrete Fourier Transform", *IEEE Transactions on Acoustics Speech and Signal Processing*, vol. 25, no. 3, pp. 235-238, (1978 Jun.).

[10] M. Abe and J. O. Smith III, "Design Criteria for Simple Sinusoidal Parameter Estimation based on Quadratic Interpolation of FFT Magnitude Peaks", 117th Audio Engineering Society Convention, San Francisco, CA, USA, (2004 Oct.), Convention paper 6256.

[11] A. J Cooper, " The electric network frequency (ENF) as and aid to authenticating a forensic digital audio recordings- An automated approach," in *Proc. AES 33rd Conf. Audio Forensics- Theory and Practice*, Denver, CO Jun. 2008, pp. 1-6.

[12] Adi Hajj-Ahmad, Ravi Grag, Min Wu, "ENF-Based Region-of-Recording Identification for Media Signals" *IEEE Transactions On Information Forensics And Security, Vol.10, No.6, Jun 2015.*

[13] P. P. Vaidyanathan, "Multirate Systems And Filter Banks". *Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.*

[14] A. Neumaier and T. Schneider, 2001: "Estimation of parameters and eigenmodes of multivariate autoregressive models". ACM Trans. Math. Softw., 27, 27C57.

[15] Grigoras, Catalin. "Applications of ENF analysis in forensic authentication of digital audio and video recordings." *Journal of the Audio Engineering Society*57.9 (2009): 643-661.

[16] Fürnkranz, Johannes. "Round robin classification." *The Journal of Machine Learning Research* 2 (2002): 721-747.
[17] Escalera, Sergio, Oriol Pujol, and Petia Radeva. "Separability of ternary codes for sparse designs of error-correcting output codes." *Pattern Recognition Letters* 30.3 (2009): 285-297.

Riyasat Ohib attained his Secondary School Certificate (S.S.C.), and Higher Secondary Certificate (H.S.C.) from Rangpur Cadet College, Rangpur. He is currently pursuing his final year of Bachelor's in Electrical and Electronic Engineering at Islamic University of Technology (IUT). His research interests lies mainly in Digital Signal Processing and Computational Electromagnetics.

Samin Yeasar Arnob has done his Secondary School Certificate (S.S.C.) from Milestone College, Dhaka and Higher Secondary Certificate (H.S.C.) from Rajuk Uttara Model College, Dhaka. Presently Samin Yeasar is pursuing his Bachelor's in Electrical and Electronic Engineering at Islamic University of Technology (IUT). He is currently involved with research in Digital Signal Processing and Photonics.

Mohammad Muhtady Muhaisin completed his Secondary School Certificate and Higher Secondary Certificate in 2010 and 2012 respectively from Sylhet Cadet College. He is a senior year student of the faculty of Electrical and Electronic Engineering at Islamic University of Technology (IUT). He is the co-author of the paper "Expedition on Mars: Timeline 2018" which was recognized and published by Mars Society in 2014. He is currently performing his thesis on Delay tolerant Network (DTN).

Tanzil Bin Hassan achieved his Secondary School Certificate (S.S.C.) from P.D.B. Secondary School, Dhaka in 2010 and completed his Higher Secondary School Certificate (HSC) from Notre Dame College, Dhaka in 2012. He is currently pursuing his final year of Bachelor's in Electrical and Electronic Engineering at Islamic University of Technology (IUT). He is co-author of the "Expedition on Mars: Timeline 2018" which was published by Mars Society in 2014. His research interests lie mainly in the field of cellular communication. His current research topic is on Femtocell in LTE based system.

Riazul Arefin is a final year student at EEE department in Islamic University of technology (IUT), Gazipur. He received his HSC degree from Rajuk Uttara Model College and SSC from Sylhet Govt. Pilot High School. His enthusiasm in research includes signal processing, photonics.

Md. Taslim Reza has got his Master's degree from Tampere University of Technology, Finland in the department of Biomedical Engineering during 2008. He also got his Bachelor degree from Islamic University of Technology, Bangladesh. He is now working in Islamic University of Technology as an assistant professor in the department of EEE. Currently he is doing his Ph. D. on medical image processing focused on ultrasound image processing used for cancer detection. Currently he is doing his Ph. D. on medical image processing focused on ultrasound image processing used for cancer detection. He is the course coordinator for the short course – 2015 arranged by EEE department, IUT.

Md. Ruhul Amin (M'10) was born in Rangpur, Bangladesh, on October 10, 1959. He received B.Sc. degree in engineering from the University of Rajshahi, Rajshahi, Bangladesh, in 1984, the M.Sc. degree in engineering from the Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, in 1987, and the Ph.D. degree in electrical and electronic engineering from Niigata University, Niigata, Japan, in 1996. He is a Professor of electrical and electronic engineering with the Islamic University of Technology, Gazipur. His current research interests include generation and application of high power microwaves, antenna theory, signal processing, and computational electromagnetics. He has authored and coauthored over 70 research papers in refereed journals and conference proceedings including the IEEE TRANSACTIONS, IET, and *Physical Review*. Dr. Amin is a recipient of the Sir Thomas Ward Memorial Medal awarded by the Institution of Engineers, India. He is a Commonwealth Fellow and a fellow of the Institution of Engineers, Bangladesh.